Semiparametric Estimation of the Canonical Permanent-Transitory Model of Earnings Dynamics*

Yingyao Hu† Robert Moffitt‡ Yuya Sasaki§

March 5, 2019

Abstract

This paper presents identification and estimation results for a flexible state space model. Our modification of the canonical model allows the permanent component to follow a unit root process and the transitory component to follow a semiparametric model of a higher-order autoregressive-moving-average (ARMA) process. Using panel data of observed earnings, we establish identification of the nonparametric joint distributions for each of the permanent and transitory components over time. We apply the identification and estimation method to the earnings dynamics of U.S. men using the Panel Survey of Income Dynamics (PSID). The results show that the marginal distributions of permanent and transitory earnings components are more dispersed, more skewed, and have fatter tails than the normal and that earnings mobility is much lower than for the normal. We also find strong evidence for the existence of higher-order ARMA processes in the transitory component, which lead to much different estimates of the distributions of and earnings mobility in the permanent component, implying that misspecification of the process for transitory earnings can affect estimated distributions of the permanent component and estimated earnings dynamics of that component. Thus our flexible model implies earnings dynamics for U.S. men different from much of the prior literature.

Keywords: earnings dynamics, semiparametric estimation, state space model

*We benefited from useful comments by Manuel Arellano, Tony Smith, seminar participants at Emory University and conference participants at the 2015 World Congress of the Econometric Society as well as those of the editors and two referees of this journal. The authors are responsible for all remaining errors.

†Johns Hopkins University, Department of Economics, Wyman Park Building 5th Floor, 3100 Wyman Park Drive, Baltimore, MD 21211

‡Johns Hopkins University, Department of Economics, Wyman Park Building 5th Floor, 3100 Wyman Park Drive, Baltimore, MD 21211

§Vanderbilt University, Department of Economics, VU Station B #351819, 2301 Vanderbilt Place, Nashville, TN 37235-1819

1 Introduction

Methods of estimating models with panel data have a long history. Those methods were first developed in the 1950s and 1960s for panel data sets of firms and of state aggregates for consumption (see Nerlove (2002) for a recounting of this period of development and for the key historical references). What we term the "canonical" model was developed in that period, consisting of a permanent component and a transitory component, distributed independently of each other. In some variants, the transitory component was assumed to follow a simple low-order ARMA process. Because of its simplicity, its intuition, and its alignment with economic theories which have permanent and transitory processes, the model has been enormously influential and has found applications in dozens of areas. Models of earnings dynamics, consumption dynamics, dynamics for firms or industries, and dynamics for individual health, student academic achievement, and other individual outcomes are just a few examples of applications.

This paper considers the identification and estimation of the canonical model under nonparametric assumptions on the unobservables. While the literature on panel data models since their development is enormous, most papers have generalized the model with additional parametric features (random walks, random growth terms, higher-order ARMAs, and other stochastic processes) and most have concerned themselves with fitting the parameters of the model only to the second moments of the data and hence fitting only the second moments of the unobservables. Our goals are to determine under what assumptions the full distribution of the unobservables in the model can be nonparametrically identified, to provide an estimator for the relevant distributions, and to provide an empirical application.

We first establish identification for our model, which is a somewhat modified version of the canonical model in several respects. For example, we allow a slightly generalized version of the common MA process, allowing it to be nonlinear; we allow the AR process to be nonstationary and to change with age; and we do not assume the shocks in each period to be i.i.d. We prove identification of the model by showing that the key unobserved elements have repeated measurements with classical measurement errors. We can, therefore, make use of the Kotlarski's identity (Kotlarski, 1967; Rao, 1992; Li and Vuong, 1998; Schennach, 2004; Bonhomme and Robin, 2010; Evdokimov, 2010) to provide closed-form identification of the distribution of the unobservables. In the identification of the generalized MA process, we rely on a recently developed result for nonlinear measurement error models (Schennach and Hu, 2013). We also provide an estimator based on deconvolution methods, which is similar to the existing estimators developed for this closed-form identification results (Li and Vuong, 1998). An advantage of this closed-form estimator is that it requires many fewer nuisance parameters than alternative semiparametric estimators.

Prior work on nonparametric identification and estimation of the canonical model and expanded versions of it include Horowitz and Markatou (1996) and Bonhomme and Robin (2010). Our paper differs from those by its approach. While the existing identification results for dynamic models with latent variables rely on a Markovian property of the dynamic structure, our paper complements the existing literature by showing the identification of a semiparametric unit-root process of a permanent state variable and a semiparametric non-Markovian process of a transitory state variable. In particular, the transitory state variable is generated by an ARMA process and does not follow a finite-order Markov process.[1] Nonparametric approaches applied to earnings dynamics models have also been developed by Geweke and Keane (2000), who allow some of the unobservables to be a mixture of normals, and by Arellano, Blundell, and Bonhomme (2017), who replace the unit root process on the permanent component with a nonparametric autoregressive function while maintaining an independence assumption for the transitory error. Our model keeps the unit root process and allow the transitory shocks to follow a semiparametric ARMA process as in the canonical models. As mentioned above, such a process of the transitory state is not Markovian and therefore can capture different dynamic structures. As for methodology, Arellano, Blundell, and Bonhomme (2017) use the results in Hu and Schennach (2008) for a general nonlinear nonclassical measurement error model with three observables. Our paper uses the Kotlarski's identity (Kotlarski, 1967; Rao, 1992; Li and Vuong, 1998; Schennach, 2004; Bonhomme and Robin, 2010; Evdokimov, 2010) and the results in Schennach and Hu (2013) for a nonlinear model with classical measurement errors when only two observables are available.

We also provide an application to the earnings dynamics of U.S. men using the Panel Study on Income Dynamics (PSID), the data set most commonly used in the literature on estimating models of individual earnings dynamics. There is a very large literature on applications to earnings dynamics models, going back to early work by Hause (1977), Lillard and Willis (1978), MaCurdy (1982), and Abowd and Card (1989), followed by many contributions including those by Horowitz and Markatou (1996), Baker (1997), Meghir and Pistaferri (2004), Guvenen (2007, 2009), Bonhomme and Robin (2010), Browning, Ejrnaes, and Alvarez (2010), Hryshko (2012), Jensen and Shore (2014), Arellano, Blundell, and Bonhomme (2017), and Botosaru and Sasaki (2018). A review of this literature, including studies which have allowed the dynamic processes to shift with calendar time, can be found in Moffitt and Zhang (2018).

Our results show that the marginal distributions of log earnings of U.S. men are nonnormal, with significant skewness and fatter tails of both the permanent and transitory components of earnings than the normal. We also find earnings dynamics very different than the normal, for our results show that the likelihood of remaining in a lower tail of the permanent earnings

[1] In fact, the AR process is a higher-order Markov process, but the MA process is not a finite-order Markov.

distribution does not fall over time as much, suggesting considerably less earnings mobility than would be found with a multivariate normality assumption. Another important finding from our empirical analysis is that the estimates of the marginal distributions as well of persistence and dynamics of permanent earnings are very sensitive to the degree of persistence in the transitory component. We find evidence for the existence of higher-order ARMA processes in the transitory component and that, with such higher-order processes, the permanent component of earnings has much less variability in marginal distributions and less mobility over time. Thus the transitory component makes a much stronger relative contribution to the marginal earnings distributions and to earnings mobility than in much of the prior literature, which often allows much less persistence in the transitory component. Finally, we consider earnings dynamics in subsamples of men with strong labor force attachment and of married men (both subsamples have been studied in the literature), finding both subsamples to have lower variances of permanent and transitory shocks than for the full population but also more earnings mobility than that population.

The rest of the paper is organized as follows. Section 2 introduces a generalized semiparametric canonical model of earnings dynamics. Section 3 presents an informal illustration of the identification strategy. Section 4 presents the formal identification results. Section 5 proposes estimators. Section 6 presents the empirical application. Section 7 concludes. Mathematical proofs, large sample properties, and additional empirical results are found in the appendix and the supplementary material.

2 The Semiparametric State Space Model

We consider the following setup of a semiparametric state space model. The measurement Y_t in time t is decomposed into two independent components:

$$Y_t = U_t + V_t. \tag{2.1}$$

The first one, U_t is the permanent state which follows the unit root process:

$$U_t = U_{t-1} + \eta_t \tag{2.2}$$

with innovation η_t. The second one, V_t is the transitory state which follows the ARMA(p,q) process:

$$V_t = \rho_{t,1} V_{t-1} + \rho_{t,2} V_{t-2} \cdots + \rho_{t,p} V_{t-p} + G_t \left(\varepsilon_t, \varepsilon_{t-1}, \cdots, \varepsilon_{t-q}\right). \tag{2.3}$$

For a short-hand notation, we write the vector of the AR coefficients by $\rho_t = (\rho_{t,1}, \cdots, \rho_{t,p})'$. Note that the time effect is the source of non-stationarity in this model both through the time-varying ARMA specifications (i.e, ρ_t and G_t) and through arbitrary time variations in the distributions of the primitives (i.e., η_t and ε_t). Because of the nonparametric specification of these time-varying distributions of the primitives, the time effect may appear in higher-order moments as well as in the first moment e.g., as commonly introduced by additive time effects in (2.1), as is common in applications. In contrast to much of the literature, we allow arbitrarily high-order ARMA processes and this will be a major feature of our empirical application in Section 6.

Our first goal in this paper is the identification of the nonparametric distributions of U_t, V_t, η_t, and ε_t as well as the function G_t and the AR parameters ρ_t in this state space model. The following example illustrates an application of this general framework to a semiparametric model of earnings dynamics.

Example 1 (The Model of Earnings Dynamics)**.** *One application is the model of earnings dynamics, where the measurement Y_t is the observed earnings at age t, the permanent state U_t is the permanent component of earnings at age t, the innovation η_t is the permanent shock at age t, and the transitory state V_t is the transitory component of earnings at age t.*

3 An Illustration of the Identification Strategy

For an illustration, we focus on the model where the permanent state follows the unit root process and the transitory state follows an ARMA(1,1) process. The general identification results will follow in Section 4. In a random sample, we observe the joint distribution of Y_t for periods $t = 1, 2, \cdots, T$. While we keep the parts (2.1) and (2.2) of the general model, the ARMA part (2.3) simplifies to

$$V_t = \rho_t V_{t-1} + G_t (\varepsilon_t, \varepsilon_{t-1}) \tag{3.1}$$

in the current section. The unknown coefficient ρ_t and the unknown function G_t may be time-varying. Furthermore, we do not require a parametric or semiparametric specification of G_t. We assume the following independence condition.

Assumption 1. *(i) The random variables $\eta_T, \cdots, \eta_1$, U_0, $\varepsilon_T, \cdots, \varepsilon_1$, and the random vector (ε_0, V_0) are mutually independent, i.e.,*

$$f(\eta_T, \cdots, \eta_1, U_0, \varepsilon_T, \cdots, \varepsilon_1, \varepsilon_0, V_0) = f(\eta_T) \cdots f(\eta_1) f(U_0) f(\varepsilon_T) \cdots f(\varepsilon_1) f(\varepsilon_0, V_0).$$

(ii) $(\eta_T, \cdots, \eta_1, U_0, V_0)$ have zero means and $E[G_t(\varepsilon_t, \varepsilon_{t-1})] = 0$ for $t \in \{1, \cdots, T\}$.

This assumption implies that process $\{U_t\}$ is independent of process $\{V_t\}$. We leave the marginal distributions of η_t and ε_t unspecified and allow them to vary arbitrarily with t. In this setup, we are interested in identification of the nonparametric distributions of the primitives ε_t and η_t, the structures ρ_t and G_t, and the nonparametric distributions of the components U_t and V_t. Our identification strategy is illustrated below in four steps.

3.1 Step 1: Identification of f_{V_t}

Consider the first difference:

$$\Delta Y_{t+1} = Y_{t+1} - Y_t = (U_{t+1} - U_t) + (V_{t+1} - V_t) \;=\; \eta_{t+1} + V_{t+1} - V_t. \tag{3.2}$$

This equation implies that we may replace V_{t+1} by V_t, η_{t+1} and ΔY_{t+1} as

$$V_{t+1} = V_t - \eta_{t+1} + \Delta Y_{t+1}. \tag{3.3}$$

Consider the following first difference for the next time period:

$$\begin{aligned}\Delta Y_{t+2} \;&=\; Y_{t+2} - Y_{t+1} \qquad (3.4)\\ &=\; \eta_{t+2} + V_{t+2} - V_{t+1} \;=\; (\rho_{t+2} - 1)\, V_{t+1} + G_{t+2}\left(\varepsilon_{t+2}, \varepsilon_{t+1}\right) + \eta_{t+2}.\end{aligned}$$

Replacing V_{t+1} by the expression in equation (3.3), we obtain

$$\frac{\Delta Y_{t+2}}{\rho_{t+2} - 1} - \Delta Y_{t+1} \;=\; V_t + \frac{G_{t+2}\left(\varepsilon_{t+2}, \varepsilon_{t+1}\right) + \eta_{t+2}}{\rho_{t+2} - 1} - \eta_{t+1} \;\equiv\; V_t + e_{t+1}. \tag{3.5}$$

With the pair of equations (2.1) and (3.5), we obtain two measurements, $\frac{\Delta Y_{t+2}}{\rho_{t+2}-1} - \Delta Y_{t+1}$ and Y_t up to an unknown scalar parameter ρ_{t+2}, of the latent variable V_t with classical measurement errors, U_t and e_{t+1}, satisfying the mutual independence among V_t, U_t and e_{t+1}. By Kotlarski's identity (under regularity conditions to be formally stated as Assumptions 9 and 10 in Section 4 for the general setup), the distribution of V_t is identified up to the unknown scalar parameter ρ_{t+2} as

$$\begin{aligned} f_{V_t}(v) \;&=\; \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i\tau v}\phi_{V_t}(\tau)d\tau, \qquad \text{where } i = \sqrt{-1} \qquad (3.6)\\ \phi_{V_t}(\tau) \;&=\; \exp\left[\int_0^{\tau} \frac{iE\left[\left(\frac{\Delta Y_{t+2}}{\rho_{t+2}-1} - \Delta Y_{t+1}\right)\exp\left(isY_t\right)\right]}{E\left[\exp\left(isY_t\right)\right]} ds\right].\end{aligned}$$

For the current step, a well-definition of the last identifying formula requires the following non-unit root assumption for the transitory state.

Assumption 2. $\rho_t \neq 1$ *for all* t.

3.2 Step 2: Identification of ρ_t

The previous step shows identification of f_{V_t} up to the unknown scalar parameter ρ_t. We now discuss alternative routes of identifying the AR parameter ρ_t. Combining (3.1) and (3.4), we obtain

$$\begin{aligned}\Delta Y_{t+2} &= (\rho_{t+2}-1)\,V_{t+1} + G_{t+2}(\varepsilon_{t+2},\varepsilon_{t+1}) + \eta_{t+2} \\ &= (\rho_{t+2}-1)\,(\rho_{t+1}V_t + G_{t+1}(\varepsilon_{t+1},\varepsilon_t)) + G_{t+2}(\varepsilon_{t+2},\varepsilon_{t+1}) + \eta_{t+2} \end{aligned} \tag{3.7}$$

Eliminating V_t with (3.5) yields

$$\begin{aligned}&\frac{\Delta Y_{t+2}}{(\rho_{t+2}-1)\,\rho_{t+1}} - \left(\frac{\Delta Y_{t+2}}{\rho_{t+2}-1} - \Delta Y_{t+1}\right) \\ =\;& \frac{(\rho_{t+2}-1)\,G_{t+1}(\varepsilon_{t+1},\varepsilon_t) + G_{t+2}(\varepsilon_{t+2},\varepsilon_{t+1}) + \eta_{t+2}}{(\rho_{t+2}-1)\,\rho_{t+1}} - \left(\frac{G_{t+2}(\varepsilon_{t+2},\varepsilon_{t+1}) + \eta_{t+2}}{\rho_{t+2}-1} - \eta_{t+1}\right).\end{aligned}$$

Notice that the last expression is independent of $Y_{t-1} = V_{t-1} + U_{t-1}$ under Assumption 1, and we get the moment restriction

$$cov\left(\left(\frac{1-\rho_{t+1}}{\rho_{t+1}(1-\rho_{t+2})}\Delta Y_{t+2} - \Delta Y_{t+1}\right), Y_{t-1}\right) = 0. \tag{3.8}$$

For a better view, we rewrite it as

$$\rho_{t+1}\frac{1-\rho_{t+2}}{1-\rho_{t+1}} = \frac{cov\,(\Delta Y_{t+2}, Y_{t-1})}{cov\,(\Delta Y_{t+1}, Y_{t-1})}. \tag{3.9}$$

We can see from this equation that, by imposing one restriction on the sequence $\rho_{t+1}, \rho_{t+2}, \cdots$, we can sequentially identify these AR parameters. Examples of such a restriction include

$$\begin{aligned}\rho_{t+1} &= \text{a known constant}, \quad \text{or} \\ \rho_{t+1} &= \rho_{t+2}.\end{aligned}$$

In the former case, one can recursively identify $\rho_{t+2}, \rho_{t+3}, \cdots$ by iterating (3.9). In the latter case, (3.9) directly yields the identifying formula

$$\rho_{t+1} = \frac{cov\,(\Delta Y_{t+2}, Y_{t-1})}{cov\,(\Delta Y_{t+1}, Y_{t-1})}, \tag{3.10}$$

provided that $cov\,(\Delta Y_{t+1}, Y_{t-1}) \neq 0$ and Assumption 2. We state this restriction as an assumption below.

Assumption 3. *$cov\,(\Delta Y_{t+1}, Y_{t-1}) \neq 0$ and $\rho_{t+1} = \rho_{t+2}$ for all t.*

3.3 Step 3: Identification of f_{η_t}, $f_{U_2,\cdots,U_{T-2}}$ and $f_{V_2,\cdots,V_{T-2}}$

Steps 1 and 2 identify the characteristic function ϕ_{V_t} by (3.6) for $t = 2, \cdots, T-2$. Given that U_t and V_t are independent, we identify the marginal distribution of U_t via the deconvolution:

$$\phi_{U_t} = \frac{\phi_{Y_t}}{\phi_{V_t}}. \tag{3.11}$$

Similarly and consequently, we also identify the marginal distribution of η_t by

$$\phi_{\eta_t} = \frac{\phi_{U_t}}{\phi_{U_{t-1}}}. \tag{3.12}$$

Notice that the independence between the permanent state U_{t-1} and the innovation η_t implies that

$$f_{U_t|U_{t-1}}(u_t, u_{t-1}) = f_{\eta_t}(u_t - u_{t-1}) \tag{3.13}$$

holds. Therefore, the joint distribution of $(U_2, U_3, \cdots, U_{T-2})$ is identified by

$$f_{U_2,U_3,\cdots,U_{T-2}} = f_{U_{T-2}|U_{T-3}} f_{U_{T-3}|U_{T-4}} \cdots f_{U_3|U_2} f_{U_2}. \tag{3.14}$$

Moreover, the independence between the process $\{U_t\}$ and the process $\{V_t\}$ implies

$$\phi_{Y_2,\cdots,Y_{T-2}} = \phi_{U_2,\cdots,U_{T-2}} \phi_{V_2,\cdots,V_{T-2}},$$

where $\phi_{Y_2,\cdots,Y_{T-2}}$ is the joint characteristic function of $Y_2, \cdots, Y_{T-2}$. Therefore, the joint distribution of the transitory states $(V_2, \cdots, V_{T-2})$ is also identified from the corresponding joint characteristic function

$$\phi_{V_2,\cdots,V_{T-2}} = \frac{\phi_{Y_2,\cdots,Y_{T-2}}}{\phi_{U_2,\cdots,U_{T-2}}}. \tag{3.15}$$

This step requires the following assumption.

Assumption 4. *(i)* $\phi_{U_1,\cdots,U_T}(s_1, \cdots, s_T) = E\left[\exp(is_1U_1 + \cdots + is_TU_T)\right]$ *is not equal to zero for any real* $(s_1, \cdots, s_T)$. *(ii) For each of* $(Y_1, \cdots, Y_T)$, $(U_1, \cdots, U_T)$, $(V_1, \cdots, V_T)$, η_t *and* ε_t, *the marginal and joint distributions are absolutely continuous with respect to the Lebesgue measure, and the marginal and joint characteristic functions are absolutely integrable.*

Part (i) of this assumption is the assumption of non-vanishing characteristic function as in Li and Vuong (1998) with a multivariate extension. It corresponds to the "completeness" assumption for nonparametric identification as in Hu and Schennach (2008) and Arellano, Blundell, and Bonhomme (2017) – see also D'Haultfoeuille (2011). In the univariate context, this assumption is known to be satisfied by most of the popular continuous distribution families, while counter-examples of distribution families violating this assumption are the uniform, the truncated normal, and many discrete distributions (Evdokimov and White, 2012). Similar remarks apply to multivariate distribution families, though there are not many stylized families of multivariate distributions. Particularly, the assumption is satisfied by the multivariate normal distributions.

We summarize the results as follows.

Proposition 1. *Suppose that Assumptions 1, 2, 3, and 4 hold. The joint distribution of* $(Y_1, \cdots, Y_T)$ *uniquely determines the marginal distribution of* η_t *for* $t = 3, 4, \cdots, T-2$, *the joint distribution of* $(U_2, \cdots, U_{T-2})$, *and the joint distribution of* $(V_2, \cdots, V_{T-2})$, *together with* ρ_t *for* $t = 3, 4, \cdots, T$.

3.4 Step 4: Identification of f_{ε_t} and G_t

Since G_t is arbitrarily nonparametric, we cannot identify the nonparametric distribution of ε_t in general. However, we may identify its distribution if the following restriction is imposed.[2]

Assumption 5. *The MA function* G_t *takes the form* $G_t(\varepsilon_t, \varepsilon_{t-1}) = \varepsilon_t + g_t(\varepsilon_{t-1})$ *with the location normalizations* $E[\varepsilon_t] = E[g_t(\varepsilon_{t-1})] = 0$.

Since we have identified ρ_t for $t = 3, 4, \cdots, T$ and the joint distribution $f_{V_2,\cdots,V_{T-2}}$, we identify the joint distribution of two composite random variables $(V_t - \rho_t V_{t-1})$ and $(V_{t-1} - \rho_{t-1} V_{t-2})$. These two random variables can be in turn rewritten as follows:

$$\begin{aligned} V_t - \rho_t V_{t-1} &= \varepsilon_t + g_t(\varepsilon_{t-1}) \\ V_{t-1} - \rho_{t-1} V_{t-2} &= \varepsilon_{t-1} + g_{t-1}(\varepsilon_{t-2}) \end{aligned} \tag{3.16}$$

The three shocks to the transitory states on the right-hand side are mutually independent. When the function $g_t(x) = \lambda_t x$ is linear, Reiersol (1950) shows that the coefficient λ_t is generally identified if ε_t is not normally distributed. Schennach and Hu (2013) generalize this result to nonlinear cases. We may identify the function g_t for $t = 4, \cdots, T-2$ and the marginal distribution of ε_t for $t = 3, \cdots, T-2$ using the results in Schennach and Hu (2013).

Assumption 6 (Schennach and Hu (2013))**.** *(i) The marginal characteristic functions of* ε_{t-1}, ε_t, $g_t(\varepsilon_{t-1})$, *and* $g_{t-1}(\varepsilon_{t-2})$ *do not vanish on the real line. (ii) The density function* $f_{\varepsilon_{t-1}}$ *of* ε_{t-1} *exists and is uniformly bounded. (iii)* g_t *is continuously differentiable, strictly monotone, and is not exactly of the form* $g_t(\varepsilon_{t-1}) = a + b\ln(e^{c\varepsilon_{t-1}} + d)$ *for* $a, b, c, d \in \mathbb{R}$.

[2] This assumption is testable via the identification result of Hu, Schennach, and Shiu (2018, Theorem 2.1), where they identify a repeated measurement model with one measurement entailing an additively separable model and the other measurement entailing a nonseparable model.

This assumption states Assumptions 1–6 and an additional condition of Theorem 1 in Schennach and Hu (2013) in terms of our notation. (The notations in Schennach and Hu (2013) and our notations are reconciled by $y := V_t - \rho_t V_{t-1}$, $x := V_{t-1} - \rho_{t-1} V_{t-2}$, $x^* := \varepsilon_{t-1}$, $\Delta y := \varepsilon_t$, $\Delta x := g_{t-1}(\varepsilon_{t-2})$ and $g := g_t$.) The first part of Assumption 1 in Schennach and Hu (2013) is implied by our Assumption 1 (ii), and hence is not included in our Assumption 6. Likewise, the second part of Assumption 1 in Schennach and Hu (2013) is implied by our Assumption 5, and hence is not included in Assumption 6. Part (i) is similar to Assumption 4 (i). As discussed earlier, it corresponds to the "completeness" assumption for nonparametric identification (D'Haultfoeuille, 2011). This assumption is known to be satisfied by most of the popular continuous distribution families, while counter-examples of distribution families violating this assumption are the uniform, the truncated normal, and many discrete distributions (Evdokimov and White, 2012). Part (ii) of the assumption is also satisfied by most of the popular continuous distribution families, with the chi-square distribution of one degree of freedom being a major counter-example. Part (iii) is a set of requirement for the function g_t in the MA decomposition.

Proposition 2. *Suppose that Assumption 5 and 6, in addition to the assumptions in Proposition 1, are satisfied. The joint distribution of* $(Y_1, \cdots, Y_T)$ *uniquely determines the marginal distribution of* ε_t *and the MA function* G_t.

This result guarantees nonparametric identification but the identification is not constructive and therefore a plug-in estimator is not available. A closed-form estimator is available at the cost of further assuming the linear MA structure as in Reiersol (1950):

$$g_t(x) = \lambda_t x.$$

In this case, (3.16) simplifies to the classical repeated measurement model:

$$\begin{aligned} V_t - \rho_t V_{t-1} &= \varepsilon_t + \lambda_t \varepsilon_{t-1} \\ V_{t-1} - \rho_{t-1} V_{t-2} &= \varepsilon_{t-1} + \lambda_{t-1} \varepsilon_{t-2} \end{aligned}$$

Therefore, we may use Kotlarski's identity to obtain the closed-form identifying formula

$$\begin{aligned} f_{\varepsilon_t}(x) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\tau x} \phi_{\varepsilon_t}(\tau) d\tau, \qquad \text{where} \\ \phi_{\varepsilon_t}(\tau) &= \exp\left[\int_0^\tau \frac{iE\left[\left(\frac{V_{t+1} - \rho_{t+1} V_t}{\lambda_{t+1}}\right) \exp\left(is\left(V_t - \rho_t V_{t-1}\right)\right)\right]}{E\left[\exp\left(V_t - \rho_t V_{t-1}\right)\right]} ds \right] \end{aligned} \tag{3.17}$$

where the expectations can be computed using the closed-form identifying formula (3.15) for the joint distribution of (V_{t-1}, V_t, V_{t+1}) obtained in the previous step.

To compute the closed form (3.17) it remains to identify the unknown scalar λ_{t+1}. We can find the moment restrictions

$$\begin{aligned} var\left(\varepsilon_{t+1}\right)+\lambda_{t+1}^{2} var\left(\varepsilon_{t}\right) &= var\left(V_{t+1}-\rho_{t+1} V_{t}\right) \\ \lambda_{t+1} var\left(\varepsilon_{t}\right) &= cov\left(V_{t}-\rho_{t} V_{t-1}, V_{t+1}-\rho_{t+1} V_{t}\right) \\ \lambda_{t+2} var\left(\varepsilon_{t+1}\right) &= cov\left(V_{t+1}-\rho_{t+1} V_{t}, V_{t+2}-\rho_{t+2} V_{t+1}\right) \end{aligned} \tag{3.18}$$

where the values on the right-hand sides can be computed again using the closed-form identifying formula (3.15) for the joint distribution of $(V_{t-1}, V_t, V_{t+1}, V_{t+2})$ obtained in the previous step. The left-hand sides contain four unknowns, $var(\varepsilon_t)$, $var(\varepsilon_{t+1})$, λ_{t+1} and λ_{t+2}. Therefore, one restriction is necessary for identification of λ_{t+1} using the above three equations.

4 General Identification Results

4.1 Nonparametric Identification of the Distributions

We now return to the general model (2.1), (2.2) and (2.3). Consider

$$\begin{aligned} Y_t &= U_t + V_t. \\ U_t &= U_{t-1} + \eta_t, \\ V_t &= \rho_{t,1} V_{t-1} + \rho_{t,2} V_{t-2} \cdots + \rho_{t,p} V_{t-p} + G_t\left(\varepsilon_t, \varepsilon_{t-1}, \cdots, \varepsilon_{t-q}\right) \end{aligned}$$

for $t = 1, 2, \cdots, T$, and the following independence and zero mean conditions.

Assumption 7 (Serial Independence and Zero Mean). *(i) The random variables $\eta_T, \cdots, \eta_1$, U_0, $\varepsilon_T, \cdots, \varepsilon_1$, and the random vector $(\varepsilon_0, \cdots, \varepsilon_{1-q}, V_0, \cdots, V_{1-p})$ are mutually independent, i.e., $f\left(\eta_T, \cdots, \eta_1, U_0, \varepsilon_T, \cdots, \varepsilon_1, \varepsilon_0, \cdots, \varepsilon_{1-q}, V_0, \cdots, V_{1-p}\right) = f\left(\eta_T\right) \cdots f(\eta_1)\ f(U_0)\ f(\varepsilon_T) \cdots f(\varepsilon_1)$ $f\left(\varepsilon_0, \cdots, \varepsilon_{1-q}, V_0, \cdots, V_{1-p}\right)$. (ii) The random vector $(\eta_T, \cdots, \eta_1, U_0, V_0, \cdots, V_{1-p})$ has zero mean and $E\left[G_t\left(\varepsilon_t, \varepsilon_{t-1}, \cdots, \varepsilon_{t-q}\right)\right] = 0$ for $t \in \{1, \cdots, T\}$.*

The goal is to derive non-parametric identification of the joint distribution of $(U_t, \cdots, U_{t+\tau})$ and the joint distribution of $(V_t, \cdots, V_{t+\tau})$. In the current subsection, we derive the identification results up to finite-dimensional AR(p) parameters $\rho_{t+q+1}, \cdots, \rho_{t+\tau+q+1}$, leaving their identification for Section 4.2. To simplify the writings, we introduce the following random and deterministic functions, μ^Y_{t+q+1}, $\nu^{\eta,\varepsilon}_{t+q+1}$ and κ, of the AR parameters ρ_{t+q+1}.

$$\mu^Y_{t+q+1}(\rho_{t+q+1}) = [Y_{t+q+1} - Y_t] - \sum_{p'=1}^{p} \rho_{t+q+1,p'} [Y_{t+q+1-p'} - Y_t]$$

$$\nu^{\eta,\varepsilon}_{t+q+1}(\rho_{t+q+1}) = \sum_{\tau'=1}^{q} \eta_{t+\tau'} - \sum_{p'=1}^{p} \sum_{\tau'=1}^{q+1-p'} \rho_{t+q+1,p'}\eta_{t+\tau'} + G_{t+q+1}(\varepsilon_{t+q+1}, \varepsilon_{t+q}, \cdots, \varepsilon_{t+1}) + \eta_{t+q+1}$$

$$\kappa(\rho_{t+q+1}) = \sum_{p'=1}^{p} \rho_{t+q+1,p'} - 1$$

The first line defines a random function μ^Y_{t+q+1} of ρ_{t+q+1} which is observed by econometricians (up to the finite dimensional AR parameters), using $\max\{p+1, q+2\}$ periods of panel data $(Y_{\min\{t,t+q+1-p\}}, \cdots, Y_{t+q+1})$. The second line defines a random function $\nu^{\eta,\varepsilon}_{t+q+1}$ of ρ_{t+q+1} which is not observed by econometricians. The third line defines a deterministic function κ of ρ_{t+q+1}. We derive identifying formulas that involve this κ function in denominators, and we make the following assumption to make sense of such identifying formulas.

Assumption 8 (AR(p) Restriction)**.** $\sum_{p'=1}^{p} \rho_{t,p'} \neq 1$ *for each* t.

Note that Assumption 8 guarantees $\kappa(\rho_{t+q+1}) \neq 0$. The following lemma, which follows from arithmetic operations using our model (2.1), (2.2) and (2.3), provides a relationship among the three random and deterministic functions, μ^Y_{t+q+1}, $\nu^{\eta,\varepsilon}_{t+q+1}$ and κ.

Lemma 1 (Restriction for V_t)**.** *If Assumption 8 is satisfied for the state space model (2.1), (2.2) and (2.3), then the following restriction holds:*

$$\frac{\mu^Y_{t+q+1}(\rho_{t+q+1})}{\kappa(\rho_{t+q+1})} = V_t + \frac{\nu^{\eta,\varepsilon}_{t+q+1}(\rho_{t+q+1})}{\kappa(\rho_{t+q+1})}. \tag{4.1}$$

The role of this auxiliary lemma is to construct repeated observations for V_t. Specifically, combining (2.1) and (4.1), we obtain the system

$$\begin{aligned} Y_t &= V_t + U_t \\ \frac{\mu^Y_{t+q+1}(\rho_{t+q+1})}{\kappa(\rho_{t+q+1})} &= V_t + \frac{\nu^{\eta,\varepsilon}_{t+q+1}(\rho_{t+q+1})}{\kappa(\rho_{t+q+1})} \end{aligned}$$

where the left-hand side of each equation is observed (up to finite dimensional parameters ρ_{t+q+1}), the first term on the right-hand side is the common factor V_t, and the second term on the right-hand side is an error. Thus, under the assumptions to be listed below, the Kotlarski's (1967) identity allows us to identify the marginal distributions of U_t and V_t as in Li and Vuong (1998). Once U_{t-1} and U_t are identified, we can in turn use the relation (2.2) to identify the marginal distribution of η_t by the deconvolution. To formally obtain these results, we note with the notation $\mathcal{I}_t = \sigma(U_t, V_t, U_{t-1}, V_{t-1}, \cdots, U_1, V_1)$ for the information available at time t that the following mean independence conditions hold under Assumption 7.

$$\mathrm{E}\left[G_\tau\left(\varepsilon_\tau, \varepsilon_{\tau-1}, \cdots, \varepsilon_{\tau-q}\right) \mid \mathcal{I}_t\right] = 0 \qquad \text{whenever } \tau - q > t. \tag{4.2}$$

$$\mathrm{E}[\eta_\tau \mid \mathcal{I}_t] = 0 \qquad \text{whenever } \tau > t. \tag{4.3}$$

The moment condition (4.2) follows from the definition of the model where the moving average is of order q. Equation (4.3) implies that the permanent state U_t follows the Martingale process, an assumption which is commonly made in the canonical models.

The marginal characteristic function ϕ_X of a random variable X is defined by $\phi_X(s) = \mathrm{E}[e^{isX}]$. The marginal characteristic functions of Y_t, U_t and V_t are denoted by ϕ_{Y_t}, ϕ_{U_t} and ϕ_{V_t}, respectively. The joint characteristic function ϕ_{X_1,X_2} of a random vector (X_1, X_2) is defined by $\phi_{X_1,X_2}(s_1, s_2) = \mathrm{E}[e^{is_1X_1+is_2X_2}]$. The joint characteristic function of (Y_τ, Y_t) is denoted by ϕ_{Y_τ,Y_t}. With these notations, we make the following regularity assumptions.

Assumption 9 (Regularity Conditions)**.** *(i) $E[Y_t]$ exists for each t. (ii) The characteristic functions ϕ_{U_t} and ϕ_{V_t} do not vanish on the real line for each t. (iii) The characteristic function ϕ_{V_t} is continuous for each t.*

Part (i) is sufficient for the existence of the moment $\mathrm{E}\left[\mu^Y_{t+q+1}(\rho_{t+q+1})\ e^{is'Y_t}\right]$ which shows up in our closed-form identifying formulas. Part (ii) is to guarantee that the denominator of the identifying formula is non-zero along with Assumption 8. It is satisfied by the major distribution families, including the normal, chi-squared, Cauchy, gamma, and exponential distributions. Part (iii) is used to recover the characteristic function of V_t from an ordinary differential equation with an initial value. By the aforementioned deconvolution approaches, we identify the marginal distributions of U_t, V_t and η_t up to the finite-dimensional parameters ρ_{t+q+1} as follows.

Lemma 2 (Identification of the Marginal Characteristic Functions)**.** *If Assumptions 7, 8, and 9 are satisfied for the state space model (2.1), (2.2) and (2.3), then ϕ_{V_t} is identified up to the finite-dimensional parameters ρ_{t+q+1} by*

$$\phi_{V_t}(s;\rho_{t+q+1}) \quad = \quad \exp\left[\int_0^s \frac{i\ \ E\left[\mu^Y_{t+q+1}(\rho_{t+q+1})\ e^{is'Y_t}\right]}{\kappa(\rho_{t+q+1})\, E\left[e^{is'Y_t}\right]} ds'\right] \tag{4.4}$$

using $\max\{p+1, q+2\}$ *periods of panel data* $(Y_{\min\{t,t+q-p+1\}}, \cdots, Y_{t+q+1})$*. Likewise, ϕ_{U_t} is identified up to the finite-dimensional parameters ρ_{t+q+1} by*

$$\phi_{U_t}(s;\rho_{t+q+1}) \quad = \quad \frac{E\left[e^{isY_t}\right]}{\phi_{V_t}(s;\rho_{t+q+1})} \tag{4.5}$$

using $\max\{p+1, q+2\}$ *periods of panel data* $(Y_{\min\{t,t+q-p+1\}}, \cdots, Y_{t+q+1})$. *Furthermore,* ϕ_{η_t} *is identified up to the finite-dimensional parameters* ρ_{t+q} *and* ρ_{t+q+1} *by*

$$\phi_{\eta_t}(s; \rho_{t+q}, \rho_{t+q+1}) \; = \; \frac{\phi_{U_t}(s; \rho_{t+q+1})}{\phi_{U_{t-1}}(s; \rho_{t+q})} \tag{4.6}$$

using $\max\{p+2, q+3\}$ *periods of panel data* $(Y_{\min\{t-1,t+q-p\}}, \cdots, Y_{t+q+1})$.

Whereas this lemma provides the identification of the marginal distributions, our goal is to identify the joint distributions. To this goal, we note that the following joint independence restrictions hold under Assumption 7 (i).

$$(U_t, \cdots, U_{t+\tau}) \perp\!\!\!\perp (V_t, \cdots, V_{t+\tau}) \tag{4.7}$$

$$\eta_{t+t'} \perp\!\!\!\perp (U_t, \cdots, U_{t+t'-1}) \text{ for any } t' \in \{1, 2, \cdots\}. \tag{4.8}$$

Furthermore, the following regularity assumptions are made for density representation of distributions and for the purpose of applying the Fourier transform.

Assumption 10 (Regularity)**.** *(i) The distribution of* U_t *is absolutely continuous with respect to the Lebesgue measure. (ii)* ϕ_{U_t} *is absolutely integrable. (iii) The distribution of* η_t *is absolutely continuous with respect to the Lebesgue measure. (iv)* ϕ_{η_t} *is absolutely integrable. (v)* $\phi_{V_t,\cdots,V_{t+\tau}}$ *is absolutely integrable. (vi)* $\phi_{U_t,\cdots,U_{t+\tau}}(s_1, \cdots, s_{t+\tau}) \neq 0$ *for all* $(s_1, \cdots, s_{t+\tau}) \in \mathbb{R}^{\tau+1}$.

Parts (i) and (iii) allow the density representation of the respective probability distributions. Parts (ii), (iv) and (v) guarantee that we can recover the density functions from the respective characteristic functions. Part (vi) plays a similar role to Assumption 9 (ii). Under these conditions, we identify the joint density of $(U_t, \cdots, U_{t+\tau})$ and the joint density of $(V_t, \cdots, V_{t+\tau})$ as follows.

Theorem 1 (Identification of the Joint Density Functions)**.** *If Assumptions 7, 8, 9 and 10 are satisfied for the state space model (2.1), (2.2) and (2.3), then* $f_{U_t,\cdots,U_{t+\tau}}$ *is identified up to the finite-dimensional parameters* $\rho_{t+q+1}, \cdots, \rho_{t+\tau+q+1}$ *by*

$$f_{U_t,\cdots,U_{t+\tau}}(u_t, \cdots, u_{t+\tau}; \rho_{t+q+1}, \cdots, \rho_{t+\tau+q+1}) = \tag{4.9}$$
$$\frac{1}{2\pi} \int e^{-isu_t} \; \phi_{U_t}(s; \rho_{t+q+1}) ds \prod_{\tau'=1}^{\tau} \left[\frac{1}{2\pi} \int e^{-is(u_{t+\tau'} - u_{t+\tau'-1})} \; \phi_{\eta_{t+\tau'}}(s; \rho_{t+\tau'+q}, \rho_{t+\tau'+q+1}) ds \right]$$

using $\max\{p+\tau+1, q+\tau+2\}$ *periods of panel data* $(Y_{\min\{t,t+q-p+1\}}, \cdots, Y_{t+\tau+q+1})$*, where* $\phi_{U_t}(s; \rho_{t+q+1})$ *and* $\phi_{\eta_t}(s; \rho_{t+q}, \rho_{t+q+1})$ *are given by (4.5) and (4.6), respectively. In addition,* $f_{V_t,\cdots,V_{t+\tau}}$ *is identified up to the finite-dimensional parameters* $\rho_{t+q+1}, \cdots, \rho_{t+\tau+q+1}$ *by*

$$f_{V_t,\cdots,V_{t+\tau}}(v_t, \cdots, v_{t+\tau}; \rho_{t+q+1}, \cdots, \rho_{t+\tau+q+1}) = \frac{1}{(2\pi)^{\tau+1}} \int \cdots \int \tag{4.10}$$
$$\frac{E\left[\prod_{\tau'=0}^{\tau} e^{is_{t+\tau'}(Y_{t+\tau'} - v_{t+\tau'})}\right]}{\int \cdots \int \prod_{\tau'=0}^{\tau} e^{is_{t+\tau'} u_{t+\tau'}} f_{U_t,\cdots,U_{t+\tau}}(u_t, \cdots, u_{t+\tau}; \rho_{t+q+1}, \cdots, \rho_{t+\tau+q+1}) du_t \cdots du_{t+\tau}} ds_t \cdots ds_{t+\tau}$$

using $\max\{p+\tau+1, q+\tau+2\}$ *periods of panel data* $(Y_{\min\{t,t+q-p+1\}}, \cdots, Y_{t+\tau+q+1})$*, where* $f_{U_t,\cdots,U_{t+\tau}}(u_t, \cdots, u_{t+\tau}; \rho_{t+q+1}, \cdots, \rho_{t+\tau+q+1})$ *is given in (4.9).*

In this theorem, (4.9) provides a closed-form identifying formula for the joint density of the permanent states $(U_t, \cdots, U_{t+\tau})$ for $\tau+1$ periods. Likewise, (4.10) provides a closed-form identifying formula for the joint density of the transitory states $(V_t, \cdots, V_{t+\tau})$ for $\tau + 1$ periods. For the both results, we need $\max\{p+\tau+1, q+\tau+2\}$ periods of panel data $(Y_{\min\{t,t+q-p+1\}}, \cdots, Y_{t+\tau+q+1})$ of measurements.

Remark 1. *It is important to observe that our general identification results for the joint density functions do not require a parametric specification of the MA part. Specifically, the identification formulae do not involve the MA parameters even if we would impose a parametric specification. In other words, the identification formulae will remain the same even if we imposed a parametric MA specification.*

4.2 Identification of the AR Parameters

The previous subsection derives the nonparametric identification of the marginal and joint distributions of the permanent state and the transitory state. These non-parametric identification results, however, assume that the finite-dimensional AR parameters $\rho_{t+q+1}, \cdots, \rho_{t+\tau+q+1}$ are already known. The current subsection explores alternative routes of identifying these remaining parameters. As a useful device to this goal, we develop the following moment equality that holds under the zero conditional mean restrictions in Assumption 7.

Proposition 3 (Moment Equality)**.** *If Assumptions 7 and 8 are satisfied for the state space model (2.1), (2.2) and (2.3), then the following moment equality holds:*

$$E\left[\kappa(\rho_t)\mu^Y_{t+1}(\rho_{t+1}) - \kappa(\rho_{t+1})\mu^Y_t(\rho_t) - \kappa(\rho_t)\kappa(\rho_{t+1})\left(Y_{t-q} - Y_{t-q-1}\right)\middle| \mathcal{I}_{t-q-1}\right] = 0. \tag{4.11}$$

The following three examples illustrate normalizing restrictions on the AR parameters ρ_t to identify them using (4.11). The first example suggests to impose the time-invariance in the AR parameters, i.e., $\rho_t = \rho_{t+1}$. The second example suggests that the initial AR parameters ρ_t are known values, and the succeeding AR parameters ρ_{t+1}, ρ_{t+2}, $\cdots$ are inductively identified. The third example suggests to impose a parametric life-cycle restriction on the AR parameters, i.e., $\rho_t = h(t, \theta)$, and to determine θ via the moment restriction (4.11).

Example 2 (Normalizing Restriction I)**.** *One normalizing restriction is the time-invariant AR process, i.e.,* $\rho_t = \rho_{t+1}$*. In this case, with* $\bar\rho := \rho_t = \rho_{t+1}$*, the moment equality (4.11) reduces to*

$$E\left[\mu^Y_{t+1}(\bar\rho) - \mu^Y_t(\bar\rho) - \kappa(\bar\rho)\left(Y_{t-q} - Y_{t-q-1}\right)\middle| \mathcal{I}_{t-q-1}\right] = 0.$$

Using $(Y_{t-q-1}, \cdots, Y_{t-q-p})$ *as instruments yields the closed-form identifying formula*

$$\bar{\rho} = E[(Y_{t-q-1}, \cdots, Y_{t-q-p})'(\Delta_{t,1}, \cdots, \Delta_{t,p})]^{-1} \, E[(Y_{t-q-1}, \cdots, Y_{t-q-p})'\Delta_{t,0}],$$

where $\Delta_{t,p'} := Y_{t+1-p'} - Y_{t-q} - Y_{t-p'} + Y_{t-q-1} + Y_{t-q} - Y_{t-q-1}$ *for each* $p' \in \{0, 1, \cdots, p\}$. □

Example 3 (Normalizing Restriction II)**.** *Another normalizing restriction is to set the initial AR parameters* ρ_t *to a p-vector of known values,* $\rho_t = \bar{\rho}$*. In this case, the moment equality (4.11) can be applied upward-inductively to recover* ρ_{t+1}*,* ρ_{t+2}*, and so on. Specifically, given* ρ_t*, we can identify* ρ_{t+1} *by the closed-form formula*

$$\rho_{t+1} = E[(Y_{t-q-1}, \cdots, Y_{t-q-p})'(\Lambda_{t,1}(\rho_t), \cdots, \Lambda_{t,p}(\rho_t))]^{-1} \, E[(Y_{t-q-1}, \cdots, Y_{t-q-p})'\Lambda_{t,0}(\rho_t)],$$

where $\Lambda_{t,p'}(\rho_t) := \kappa(\rho_t)\,[Y_{t+1-p'} - Y_{t-q-1}] + \mu_t^Y(\rho_t)$ *for each* $p' \in \{0, 1, \cdots, p\}$. □

Example 4 (Parametric Life-Cycle Restriction)**.** *We may specify the sequence of the AR parameters* ρ_t *as a parametric function of* t*, i.e.,* $\rho_t = h(t, \theta)$*. We may then use the moment equality (4.11) to construct the moment function*

$$\begin{aligned} g(\theta) := E\big[(Y_{t-q-1}, \cdots, Y_{t-q-p})' \big\{&\kappa(h(t,\theta))\mu_{t+1}^Y(h(t+1,\theta)) - \kappa(h(t+1,\theta))\mu_t^Y(h(t,\theta)) \\ &-\kappa(h(t,\theta))\kappa(h(t+1,\theta))\,(Y_{t-q} - Y_{t-q-1})\big\}\big] \end{aligned}$$

for a GMM estimation of θ*, and thus for* $\rho_t = h(t, \theta)$ *for all* t*.* □

For generality to encompass all the above examples, we state the conditions for the identification of the AR parameters as a high-level assumption below.

Assumption 11 (Identification of the AR Parameters)**.** *The moment equality (4.11) admits a unique solution* $(\rho_{t+q+1}, \cdots, \rho_{t+\tau+q+1})$*.*

The main identification result is now stated as the following corollary to Theorem 1.

Corollary 1 (Identification of the Joint Density Functions)**.** *If Assumptions 7, 8, 9, 10, and 11 are satisfied for the state space model (2.1), (2.2) and (2.3), then* $f_{U_t,\cdots,U_{t+\tau}}$ *is identified by (4.9) using* $\max\{p+\tau+1, q+\tau+2\}$ *periods of panel data* $(Y_{\min\{t,t+q-p+1\}}, \cdots, Y_{t+\tau+q+1})$*. In addition,* $f_{V_t,\cdots,V_{t+\tau}}$ *is identified by (4.10) using* $\max\{p+\tau+1, q+\tau+2\}$ *periods of panel data* $(Y_{\min\{t,t+q-p+1\}}, \cdots, Y_{t+\tau+q+1})$*.*

Remark 2. *Recall that we also show the identification of the MA structure* G_t *in Section 3.4 under the additive separability restriction (Assumption 5). Given that* G_t *is arbitrarily nonparametric and that* ε_t*, ...,* ε_{t-q} *are nonparametrically distributed, it is difficult to identify* G_t *unless some model restriction is imposed, such as the additive separability restriction (Assumption 5). A potential direction to proceed without imposing the additive separability restriction* *is to impose support restrictions as in Hu and Sasaki (2017). Under their support restriction assumption, nonseparable repeated measurement models are indeed identifiable, although it is not necessarily easy to argue that their assumption of support restrictions is satisfied in general for the application to earnings dynamics. Generalized identification of a nonseparable models with repeated measurements deserves a topic for future research.*

4.3 Relation with an Existing Approach

Our identification results are based on the deconvolution method with Kotlarski's identity (e.g., Kotlarski, 1967; Rao, 1992; Li and Vuong, 1998; Schennach, 2004; Bonhomme and Robin, 2010; Evdokimov, 2010), and are closely related to the recent econometrics literature on identification of latent process models (e.g., Arellano, Blundell, and Bonhomme, 2017) based on an operator theoretic approach (e.g., Hu and Schennach, 2008). The objectives are quite similar – one is interested in semi- or non-parametrically identifying the joint and marginal distributions of the latent components in dynamic processes. The two approaches to identification are distinct, however, and exhibit tradeoffs in terms of model flexibility and practicality.

On one hand, the framework of Hu and Schennach (2008) and Arellano, Blundell, and Bonhomme (2017) admit nonparametric and nonseparable models with greater extents of flexibility. This contrasts with the semiparametric and additive restrictions that we impose on our model. On the other hand, the approach of Hu and Schennach (2008) and Arellano, Blundell, and Bonhomme (2017) entail implicit identification without any closed-form guide to sample counterpart estimators. All the nonparametric parts of our identification results are accompanied by explicit and closed-form identifying formulas, which in turn yield closed-form analog estimators presented in the following section.

More importantly, however, our main objective is to allow for a non-Markovian transitory state process, in particular through the ARMA model. Nonparametric or semiparametric identification under this non-conventional setting seems to require partial additivity, and we hence find the approach based on Kotlarski's identity to be a natural path.

5 Estimation

Sample counterparts of the identification results yield closed-form estimators, since we derive closed-form identification for the density functions of the permanent state and the transitory state. In this section, we propose the closed-form estimators. Many details and additional results are delegated to the appendix for a concise exposition. Large sample properties are developed by extending the results of Li and Vuong (1998) – see Appendix C in the supplementary material.

The first step is to estimate the AR coefficients following the sample analog of the moment restrictions in Example 2, 3, or 4. For Example 2, the analog estimator $\widehat{\rho}$ is

$$\widehat{\rho} = \left[\sum_{j=1}^{N}(Y_{j,t-q-1}, \cdots, Y_{j,t-q-p})'(\Delta_{j,t,1}, \cdots, \Delta_{j,t,p})\right]^{-1}\left[\sum_{j=1}^{N}(Y_{j,t-q-1}, \cdots, Y_{j,t-q-p})'\Delta_{j,t,0}\right],$$

where $\Delta_{j,t,p'} := Y_{j,t+1-p'} - Y_{j,t-q} - Y_{j,t-p'} + Y_{j,t-q-1} + Y_{j,t-q} - Y_{j,t-q-1}$ for each $p' \in \{0, 1, \cdots, p\}$. We may of course extend this estimator by pooling the sums across t from $1 + p + q$ to $T - 1$. Example 3 also entails a similarly simple parametric estimator. For estimation of ρ_t under Example 4, see Appendix B.1 in the supplementary material for detailed procedures.

The second step is to estimate the marginal characteristic functions of U_t, V_t and η_t by the sample analog of the identifying formulas displayed in Lemma 2. Specifically, the analog estimators for (4.4), (4.5) and (4.6) read

$$\begin{aligned}
\widehat{\phi}_{V_t}(s;\widehat{\rho}_{t+q+1}) &= \exp\left[\int_0^s \frac{i\ \sum_{j=1}^N \mu^Y_{j,t+q+1}(\widehat{\rho}_{t+q+1})\ e^{is'Y_{j,t}}}{\kappa(\widehat{\rho}_{t+q+1})\sum_{j=1}^N e^{is'Y_{j,t}}}ds'\right] \quad \text{for } t \in \{1, \cdots, T-q-1\}, \\
\widehat{\phi}_{U_t}(s;\widehat{\rho}_{t+q+1}) &= \frac{N^{-1}\sum_{j=1}^N e^{isY_{j,t}}}{\widehat{\phi}_{V_t}(s;\widehat{\rho}_{t+q+1})} \quad \text{for } t \in \{1, \cdots, T-q-1\}, \qquad \text{and} \\
\widehat{\phi}_{\eta_t}(s;\widehat{\rho}_{t+q},\widehat{\rho}_{t+q+1}) &= \frac{\widehat{\phi}_{U_t}(s;\widehat{\rho}_{t+q+1})}{\widehat{\phi}_{U_{t-1}}(s;\widehat{\rho}_{t+q})} \quad \text{for } t \in \{2, \cdots, T-q-1\},
\end{aligned}$$

respectively, where

$$\begin{aligned}
\mu^Y_{j,t+q+1}(\widehat{\rho}_{t+q+1}) &= [Y_{j,t+q+1} - Y_{j,t}] - \sum_{p'=1}^{p}\widehat{\rho}_{t+q+1,p'}\,[Y_{j,t+q+1-p'} - Y_{j,t}] \qquad \text{and} \\
\kappa(\widehat{\rho}_{t+q+1}) &= \sum_{p'=1}^{p}\widehat{\rho}_{t+q+1,p'} - 1.
\end{aligned}$$

The final step is to estimate the density functions. Specifically, the marginal density function V_t can be estimated by the regularized Fourier transform of the second-step estimator $\widehat{\phi}_{V_t}(s;\widehat{\rho}_{t+q+1})$:

$$\widehat{f}_{V_t}(v_t) = \frac{1}{2\pi}\int e^{-isv_t}\widehat{\phi}_{V_t}(s;\widehat{\rho}_{t+q+1})\phi_K(hs)ds \quad \text{for } t \in \{1, \cdots, T-q-1\},$$

where ϕ_K denotes the Fourier transform of a suitable choice of a kernel function K, and h denotes the bandwidth parameter – we discuss ϕ_K and h in Appendix C in the supplementary material. Likewise, the marginal density functions of U_t and η_t can be estimated by the regularized Fourier transforms

$$\begin{aligned}
\widehat{f}_{U_t}(u_t) &= \frac{1}{2\pi}\int e^{-isu_t}\widehat{\phi}_{U_t}(s;\widehat{\rho}_{t+q+1})\phi_K(hs)ds \quad \text{for } t \in \{1, \cdots, T-q-1\} \qquad \text{and} \\
\widehat{f}_{\eta_t}(\eta_t) &= \frac{1}{2\pi}\int e^{-is\eta_t}\widehat{\phi}_{\eta_t}(s;\widehat{\rho}_{t+q},\widehat{\rho}_{t+q+1})\phi_K(hs)ds \quad \text{for } t \in \{2, \cdots, T-q-1\}
\end{aligned}$$

respectively.

Furthermore, we can estimate the joint density function of $(U_t, \cdots, U_{t+\tau})$ and the joint density function of $(V_t, \cdots, V_{t+\tau})$ by the sample analog of the identifying formulas displayed in Theorem 1. Specifically, the analog estimator for (4.9) reads

$$\widehat{f}_{U_t,\cdots,U_{t+\tau}}(u_t, \cdots, u_{t+\tau}; \widehat{\rho}_{t+q+1}, \cdots, \widehat{\rho}_{t+\tau+q+1}) = \frac{1}{2\pi} \int e^{-isu_t}\ \widehat{\phi}_{U_t}(s; \widehat{\rho}_{t+q+1}) \phi_K(hs) ds$$
$$\prod_{\tau'=1}^{\tau} \left[\frac{1}{2\pi} \int e^{-is(u_{t+\tau'} - u_{t+\tau'-1})}\ \widehat{\phi}_{\eta_{t+\tau'}}(s; \widehat{\rho}_{t+\tau'+q}, \widehat{\rho}_{t+\tau'+q+1}) \phi_K(hs) ds \right]$$

for $t \in \{1, \cdots, T-q-1-\tau\}$. Likewise, the sample analog estimator for (4.10) reads

$$\widehat{f}_{V_t,\cdots,V_{t+\tau}}(v_t, \cdots, v_{t+\tau}; \widehat{\rho}_{t+q+1}, \cdots, \widehat{\rho}_{t+\tau+q+1}) = \frac{1}{(2\pi)^{\tau+1}} \int \cdots \int$$
$$\frac{N^{-1} \sum_{j=1}^{N} \left[\prod_{\tau'=0}^{\tau} e^{is_{t+\tau'}(Y_{j,t+\tau'} - v_{t+\tau'})} \right] \cdot \phi_K(H_t) \cdots \phi_K(Hs_{t+\tau})}{\int \cdots \int \prod_{\tau'=0}^{\tau} e^{is_{t+\tau'} u_{t+\tau'}} \widehat{f}_{U_t,\cdots,U_{t+\tau}}(u_t, \cdots, u_{t+\tau}; \widehat{\rho}_{t+q+1}, \cdots, \widehat{\rho}_{t+\tau+q+1}) du_t \cdots du_{t+\tau}} ds_t \cdots ds_{t+\tau}$$

for $t \in \{1, \cdots, T-q-1-\tau\}$ with the multi-dimensional regularization, where H denotes the bandwidth parameter. We use this upper case notation H to distinguish it from the previous bandwidth parameter h, where their asymptotic divergence rates are different – see Appendix C in the supplementary material for details. This multivariate density estimate $\widehat{f}_{V_t,\cdots,V_{t+\tau}}$ can be also used to estimate the MA errors ε_t under an additional model restriction described in Section 3.4 – see Appendix B.2 in the supplementary material for details.

Finally, we remark that we can also use the estimated characteristic functions in turn to estimate the moments of the latent components, U_t and V_t. The estimated moments can then be used to obtain estimates of the distributional indices, such as standard deviations, skewness, and kurtosis. Specifically, we estimate the k-th moment of V_t by $i^{-k} \frac{d^k}{ds^k} \widehat{\phi}_{V_t}(s; \widehat{\rho}_{t+q+1}) \Big|_{s=0}$. Furthermore, we estimate the k-th moment of U_t by $i^{-k} \frac{d^k}{ds^k} \left[\frac{\widehat{\phi}_{Y_t}(s)}{\widehat{\phi}_{V_t}(s; \widehat{\rho}_{t+q+1})} \right]_{s=0}$ where $\widehat{\phi}_{Y_t}(s) = \frac{1}{N} \sum_{j=1}^{N} e^{isY_{j,t}}$. See Appendix B.3 in the supplementary material for the closed-form estimators for the first four moments of V_t and U_t that are needed to compute the important distributional indices including the skewness and the kurtosis. They all consist of analytic expressions written in terms of sample moments of the measurements to admit linear representations, and hence the asymptotic normality of these moment and index estimators follows in the standard way by applications of the central limit theorem and the delta method. Large sample properties are presented in Appendix C in the supplementary material.[3]

[3]Kato, Sasaki, and Ura (2018) develop a method of inference for density functions identified by Kotlarski's identity.

6 Application to Earnings Dynamics

We apply the identification and estimation methods for our semiparametric model to the case of earnings dynamics and we estimate the distributions of the error terms and parameters of the model and show the results. We also focus on three types of analyses that demonstrate the contributions of our flexible model relative to past work. First, we analyze higher-order ARMA models for the transitory effect and we study the quantitative implications of omitting higher-order components. Second, we analyze the quantitative implications of omitting life-cycle effects in the persistence parameters for life-cycle earnings dynamics. Third, we analyze the quantitative implications of imposing Gaussian distributions for the error terms.

We first describe the data that we use for our analysis in Section 6.1, and then discuss the empirical procedure in Section 6.2. Estimation results and their discussions are presented in Sections 6.3 and 6.4 with an emphasis on the above three points.

6.1 Data

We use the most commonly-used U.S. data set for earnings dynamics, the Panel Study of Income Dynamics (PSID), 1970-1996.[4] This data set has been used by Horowitz and Markatou (1996) and Bonhomme and Robin (2010) for earnings models with related econometric approaches based on deconvolution.

Our sample selection procedure is similar to those of preceding papers on earnings dynamics using the PSID – see Moffitt and Zhang (2018) for a survey. We select male individuals aged 25–55 who are recorded as household heads. Full-time students are excluded from the sample. In the first stage, we estimate a regression of log annual earnings on education, separately by year, and use the residuals Y_t to estimate the earnings dynamics model. Extreme outliers for Y_t are trimmed at the top one percent and bottom one percent, consistent with usual practice in this literature. Allowing for a unbalanced sample from the above sample selections, we obtain $T = 31$ and NT =28,436. The cross-sectional sample sizes are N =1,320 at age 30, N =1,185 at age 40, and N =984 at age 50 to list a few age groups. The top third of Table 1, labeled as the "baseline" sample, provides summary statistics of this baseline sample.

In addition to the baseline sample defined above, we also consider two subsamples that have been studied in the literature. The first is a subsample of those individuals with strong labor force attachment, defined by 40 weeks or more of work in the previous year. In the literature, a subsample based on strong labor force attachment is considered by Guvenen (2009). We examine whether the patterns of earnings dynamics are different for this subsample and for the baseline sample. The total unbalanced sample has $T = 31$ and NT =25,328 while the

[4] We do not use the data after 1996 where interviews are conducted biannually.

cross-sectional sample sizes are N =1,183 at age 30, N =1,076 at age 40, and N =864 at age 50. The middle third of Table 1, labeled as the sample of individuals with "strong labor force attachment," provides a summary statistics of this sub-sample.

The second subsample we consider is one which selects only married men. This subsample was considered by Arellano, Blundell, and Bonhomme (2017). We test whether married men have more stable earnings dynamics than for men as a whole. The unbalanced sample has $T = 31$ and NT =25,328 and the cross-sectional sample sizes are N =991 at age 30, N =993 at age 40, and N =873 at age 50. The bottom third of Table 1, labeled as the sample of "married" individuals, provides a summary statistics of this subsample.

6.2 Empirical Procedure

In the framework of Example 4 to obtain ρ_t, we set the life-cycle of AR parameters by the cubic function $\rho_t = h(t, \theta) = \theta_0 + \theta_1 t + \theta_2 t^2 + \theta_3 t^3$ in the baseline model but we also try an alternative specification as a sensitivity analysis. The auxiliary parameters $\theta = (\theta_0, \theta_1, \theta_2, \theta_3)$ and thus the AR parameters $\rho_t = h(t, \theta)$ for each $t \in \{26, \cdots, 55\}$ are estimated using the GMM – see Appendix B.1 in the supplementary material. Since our subsequent identification steps require Assumption 8, the estimation imposes this additional restriction. Specifically, we impose $\rho_t \in (0, 1)$ for all t.

For nonparametric density estimation, we use the kernel function given in the supplementary appendix. The bandwidth parameter is chosen to minimize the integrated squared errors with the reference normal distribution with the variance corresponding to the negative second derivative of the estimated characteristic function.[5] We also estimate the MA structure focusing on the simple case of ARMA(1,1) model described in Section 3.4. Details of the estimation procedure are described in Appendix B.2 in the supplementary material. Like the AR parameter ρ_t, we set the life-cycle cubic function $\lambda_t = l(t, \vartheta) = \vartheta_0 + \vartheta_1 t + \vartheta_2 t^2 + \vartheta_3 t^3$ for the MA parameter λ_t.

6.3 Results for the Baseline Sample

6.3.1 Marginal Distributions

Table 2 shows estimates of the model assuming ARMA(0,0) and the top panels of Tables 3, 4, 5, and 6 show, respectively, estimated indices of the marginal distributions of the permanent and transitory earnings at three different ages under the ARMA(1,1), ARMA(2,2), ARMA(3,3),

[5] Delaigle and Gijbels (2004) propose a number of methods to choose the bandwidth parameter for a deconvolution estimator, although our framework does not exactly fit theirs.

and ARMA(4,4) models.[6] The displayed indices are the mean, the standard deviation, the skewness, and the kurtosis. The numbers in parentheses indicate the standard errors of the respective estimates. The last column shows the p-values for the one-sided test of the null hypothesis that kurtosis is less than equal to three, against the alternative hypothesis that it is greater than three (recall that the Gaussian distribution has the kurtosis of three, and the p-valued hence indicate results of the test of sub-Gaussianity).

Several patterns appear in all tables, regardless of the ARMA order. The estimated means are very close zero uniformly across all the models but the standard deviations are greater for the permanent component than for the transitory. Also, the standard deviations of the permanent component tend to grow with age while those for the transitory component tend to decline with age. These standard deviation patterns are consistent with past evidence showing that earnings profiles tend to spread out with age but that older workers settle into more stable earnings profiles. The distributions are negatively skewed but are more skewed for the transitory component than for the permanent, and both are quadratic in age, falling from age 30 to 40 but rising from age 40 to 50. Strong evidence of kurtosis appears in almost all distributions and Gaussianity is almost always rejected at conventional confidence levels, implying distributions that are more fat-tailed than the normal.

On the other hand, several patterns differ across the ARMA orders. For example, the estimated standard deviations exhibit heterogeneous patterns across models. Specifically, the standard deviations of the permanent component of earnings tend to decrease as the order of the ARMA model increases, while the standard deviations of the transitory component of earnings tend to increase as the order of the ARMA model increases. This implies that the lower-order models, such as ARMA(0,0), erroneously impute larger portions of cross sectional variations in earnings to variations in permanent earnings. As such, omitting higher-order terms in the ARMA models produces biases in estimates of the distributions of earnings components. The estimated skewness and the estimated kurtosis exhibit similar patterns to those of the estimated standard deviations.

There is mixed evidence on the importance of time-varying, life-cycle effects in the ARMA parameters. On the one hand, the top panel of Figure 1 presents the estimated ARMA(1,1) parameters and indicates that there exist nontrivial life-cycle effects in the persistence parameters, with rising AR parameters and falling MA parameters with age. However, the lower panels of Tables 3, 4, 5, and 6 show that our estimates of the shapes of the distributions are not much affected if life cycle effects in the ARMA parameters are ignored.

The third point we focus on, as emphasized at the beginning of Section 6, concerns the

[6] In addition to the estimated distributional indices, we also illustrate estimated marginal densities of the permanent and transitory components of earnings in the supplementary material.

Gaussianity of distributions. The hypothesis that the distributions of the permanent component of earnings have a sub-Gaussian kurtosis is rejected for ages 40 and 50 at the level of 0.05 across all the model specifications. The hypothesis that the distributions of the transitory component of earnings have a sub-Gaussian kurtosis is rejected for ages 30 and 40 at the level of 0.05 across all the model specifications except under the most restrictive model, namely ARMA(0,0). From these results, we conclude that it is too restrictive to use Gaussian marginal distributions for canonical models of earnings dynamics.

To summarize the results of marginal distributions, we find the following three points. First, omitting higher-order components of the ARMA model imply different distributions both for the permanent and transitory components of earnings. Second, omitting life-cycle effects of persistence parameters does not imply different distributions either for the permanent or transitory components of earnings. Third, it is too restrictive to impose Gaussian marginal distributions for earnings dynamics models.

6.3.2 Joint Distributions and Implications for Life-Cycle Earnings Dynamics

We now turn to an analysis of life-cycle earnings dynamics which focuses on the degree of persistence in the life cycle earnings process. Persistence is often measured with impulse response functions, showing how shocks to a variable affect the mean of a variable at later dates. Given our results in the last section, we are more interested in the tails of the distributions rather than the means. We instead measure persistence by using a measure of lower tail dependence, which is the probability that earnings fall below a particular percentile point of the distribution at age t if it was below that percentile point at age $\tau < t$. Further, we focus on lower tail dependence in the permanent component and at different ages. Thus, for example, provided that the permanent component of earnings of a worker is in the bottom one percent of the distribution at age 30, what is the probability that it stays in the bottom one percent thereafter? To answer this question, we draw trajectories of the probability that the permanent component of earnings at age $30 + \Delta$ falls below the first percentile in the cross section provided that the worker had the permanent component of earnings at age 30 below the first percentile in the cross section.[7] This conditional probability is quantified by $\lambda^l_{30,t}(0.01) = P(U_t \leq F^{-1}_{U_t}(0.01) | U_{30} \leq F^{-1}_{U_{30}}(0.01))$ where $t = 30 + \Delta$ – see Section 6.2. Under the bivariate Gaussian copula, it is well known that $\lim_{q \to 0} \lambda^l_{30,t}(q) = 0$ must hold, and hence $\lambda^l_{30,t}(0.01)$ at $q = 0.01$ is supposed to be a very small probability. As such, Gaussian models have limited abilities to describe life-cycle earnings dynamics of lower tail persistence. On the other hand, our semiparametric model can allow

[7] In addition to the estimated trajectories of these conditional probabilities, we also illustrate estimated joint densities of the permanent and transitory components of earnings in the supplementary material.

$\lambda^{l}_{30,t}(0.01)$ to possibly take a high probability unlike the case of the bivariate Gaussian copula.[8]

Figures 2a and 2b display trajectories of the lower tail dependence measure $\lambda^{l}_{30,t}(0.01)$ of permanent earnings for ages $t \in \{31, ..., 50\}$ following the event of permanent earnings less than or equal to the first percentile at age 30. The solid lines represent the trajectories under our semiparametric model and the dashed lines represent those under the bivariate normal distribution. The results are displayed under each of the ARMA(0,0), ARMA(1,1), ARMA(2,2), ARMA(3,3) and ARMA(4,4) specifications with time-varying persistence parameters and with time-invariant persistence parameters. The figures thus show how lower tail dependence depends on the ARMA order, whether life cycle effects in the ARMA process are present, and the effect of imposing Gaussianity.

In all cases, the probability of remaining in the lower first percentile drops immediately at the next age and then remains, fairly stably, thereafter. But the probability drops to a very different level depending on the ARMA order, with persistence (i.e., immobility) much higher at high-order ARMAs than at low-order ARMAs (i.e., the solid lines tend to shift upward as the order increases. These results imply again that omitting higher-order components in the ARMA model can provide restrictive implications for life-cycle earnings dynamics and, specifically, lower-order ARMAs show too little persistence and too much mobility in the permanent component. For example, under more flexible higher-order ARMA models, such as ARMA(4,4), the conditional probability $\lambda^{l}_{30,t}(0.01)$ of extremely low permanent earnings remains as high as 0.9 until age 50 under our semiparametric model.

Regarding the the importance of age-varying ARMA parameters, we do not detect qualitative differences in the trajectories between the model with time-varying persistence parameters and the model with time-invariant persistence parameters. As such, we fail to find different implications of omitting life-cycle effects in the persistence parameters for life-cycle earnings dynamics through our analysis. This conclusion is consistent with our conclusion from the analysis of marginal distributions above.

Finally, regarding normality, we find that the trajectories of $\lambda^{l}_{30,t}(0.01)$ under the bivariate Gaussian distribution (dashed lines) are consistently lower than those under the flexible semiparametric models. In particular, the Gaussian trajectories appear very close to zero under more flexible higher-order ARMA models, such as the ARMA(4,4) model. This is consistent with the well-known fact that $\lim_{q \to 0} \lambda^{l}_{30,t}(q) = 0$ holds under the Gaussian copula. On the other hand, as just noted, our semiparametric model allows the life-cycle earnings dynamics to exhibit greater tail dependence than the Gaussian model can. Thus we find again that our semiparametric model gives a different answer to earnings dynamics in the tails than would a

[8]We considered measures of upper tail dependence as well but these measures were very noisy and we do not present them.

Gaussian model.

Figure 3 displays trajectories of the lower tail dependence measures, $\lambda^l_{30,t}(0.10)$ and $\lambda^l_{30,t}(0.05)$, as well as $\lambda^l_{30,t}(0.01)$, under each of the ARMA(1,1) and ARMA(4,4) models with time-varying persistent parameters, but for the lower fifth and tenth percentile points of the permanent component distribution rather than for the first. The results show that the discrepancies imputed both to the order of the ARMA model and the Gaussianity diminish as the percentile of interest increases. In other words, misspecification of any of the three types will not cause much difference on the dynamic dependence at a higher percentile, such as the tenth percentile, of the distribution.

A final issue we consider is whether our findings vary with age. For this, we examine results for age 40 instead of 30. Figure 4 displays trajectories of the lower tail dependence measure $\lambda^l_{40,t}(0.10)$, $\lambda^l_{40,t}(0.05)$, and $\lambda^l_{40,t}(0.01)$ under the ARMA(4,4) models with time-varying persistent parameters, as well as $\lambda^l_{30,t}(0.10)$, $\lambda^l_{30,t}(0.05)$, and $\lambda^l_{30,t}(0.01)$ for the purpose of comparison. The results show that lower-tail persistence is considerably lower at higher ages than at lower ages. This is a surprising result because it implies that older workers have greater upward mobility if they have very low earnings than do younger workers. However, the qualitative patterns of the implications of distributional misspecification are the same between $\lambda^l_{30,t}(q)$ and $\lambda^l_{40,t}(q)$ for $q \in \{0.01, 0.05, 0.10\}$, for the discrepancy between the two types of the lines is the largest for the first percentile and diminishes as the percentile of interest increases.

6.4 Results for Restricted Samples

As described in Section 6.1, we consider two of restricted subsamples of the baseline sample, one a subsample of workers with strong labor force attachment and one a subsample of married workers. These two subsamples yield qualitatively very similar results to each other, and we hence focus on the subsample of workers with strong labor force attachment. A complete set of results for the subsample of married workers can be found in the supplementary appendix, but we also briefly discuss results for married workers in Section 6.4.3.

6.4.1 Marginal Distributions

Tables 7, 8, 9, 10, and 11 summarize estimated indices of the marginal distributions of the permanent and transitory earnings under the ARMA(0,0), ARMA(1,1), ARMA(2,2), ARMA(3,3), and ARMA(4,4) models with time-varying persistence parameters and with time-invariant persistence parameters for the strong labor force attachment subsample.[9]

[9] In addition to the estimated distributional indices, we also illustrate semiparametrically estimated marginal densities of the permanent and transitory components of earnings in the supplementary material.

The estimated means are very close to zero uniformly across all the models, as in the baseline sample. But the standard deviations of the permanent and transitory components are almost always somewhat smaller than those in the baseline sample, consistent with the expectation that those with strong labor force attachment have more stable profiles. Similar to the results for the baseline sample, the standard deviations of the permanent component of earnings tend to decrease as the order of the ARMA model increases while the standard deviations of the transitory component of earnings tend to increase with the order of the ARMA model. Therefore, omitting higher-order terms in the ARMA specification again can cause specification biase in the component distributions. On the other hand, similar to the results for the baseline sample, we do not detect any outstanding evidence that the life-cycle effects in the persistence parameters significantly matter for marginal distributions. The model implications for the subsample discussed thus far about the mean and the standard deviation are therefore similar to those for the baseline sample except for the magnitude of the standard deviations.

The estimated distributions of the permanent component of earnings entail significantly negative skewness at ages 30 and 40 across all the model specifications. This pattern of the results for the subsample differs from that of the results for the baseline sample, where the skewness of the permanent component distribution tends to disappear as the order of the ARMA model increases. The magnitude of the skewness, however, is always smaller for the restricted sample than for the baseline sample at age 30. Thus, workers with stronger labor force attachment have less negative skewness than do other workers, although the skewness does maintain itself later into the life cycle. Also unlike the results for the baseline model, the hypothesis that the distributions of the permanent component of earnings have a sub-Gaussian kurtosis is not rejected at ages 40 and 50 at the level of 0.05 in any of the model specifications (hence less fat-tailed than the overall population). On the other hand, the hypothesis that the distributions of the transitory component of earnings have a sub-Gaussian kurtosis is still rejected at ages 30 and 40 at the level of 0.05 across all the model specifications, similar to the results for the baseline model. However, the kurtosis of the transitory component for the restricted sample is always smaller than the kurtosis for the baseline sample under all the model specifications at ages 30 and 40. Hence the subsample of workers with strong labor force attachment has less evidence of fat tails in the transitory component of earnings.

6.4.2 Joint Distributions and Implications for Life-Cycle Earnings Dynamics

Figures 5a and 5b display trajectories of the lower tail dependence measure $\lambda^l_{30,t}(0.01)$ of permanent earnings for ages $t \in \{31, ..., 50\}$ following the event of permanent earnings less than or equal to the first percentile at age 30 for the strong labor force attachment subsample. The solid lines again represent the trajectories under our semiparametric model and the dashed lines again represent those under the bivariate normal distribution. The results are displayed under each of the ARMA(0,0), ARMA(1,1), ARMA(2,2), ARMA(3,3),and ARMA(4,4) specifications with time-varying persistence parameters and with time-invariant persistence parameters.

We obtain qualitatively the same results as those we obtained for the baseline sample. Comparing the life-cycle dynamics between the baseline sample and the restricted sample, we see persistence patterns that are very close to one another in the baseline and restricted samples under the ARMA(1,1) and ARMA(2,2) models but less persistence in the restricted sample than in the baseline sample for the ARMA(3,3) and ARMA(4,4) models. Thus, at least for these measures of persistence, we find that, for higher-order ARMA models, those with strong labor force attachment are more likely to move out of their initial quantile than the full population.

Regarding the three points of the focus of our analysis emphasized at the beginning of Section 6, we once again conclude the following three points. First, under more flexible higher-order ARMA models, such as ARMA(4,4), the conditional probability $\lambda^{l}_{30,t}(0.01)$ of permanent earnings at the first percentile remains as high as 0.7 under the semiparametric model, while it stays as low as 0.2 under the bivariate Gaussian distribution. Second, we fail to find different implications of omitting life-cycle effects in the persistence parameters for life-cycle earnings dynamics through our analysis. Third, our more flexible semiparametric model gives different answers to questions regarding life-cycle dynamics of earnings than would a Gaussian model. These three points of the conclusion are the same as for the case of the baseline sample.

6.4.3 Married Workers

In addition to the subsample of workers with strong labor force attachment analyzed above, we also estimate the model for the subsample of married workers. All the results look similar to those presented above for the subsample of workers with strong labor force attachment – see supplementary appendix for a complete set of results. Therefore, we do not repeat discussions of the results for this subsample. However, two remarks on the results are in order. First, the standard deviations of both the permanent and transitory component distributions are smaller for this subsample than for the baseline sample. Married men, therefore, have more stable earnings profiles than for men as a whole. Furthermore, married men are more likely to exit the bottom 1 percentile than was the case for men as a whole. This finding echoes that for men with strong labor force attachment referred to above.

7 Conclusions

In this paper we have investigated identification and estimation of flexible state space models. In our version of the canonical model, the permanent state U_t follows a unit root process and the transitory transitory state V_t follows a semiparametric model of ARMA(p,q) process. Using panel data of measurements Y_t, we establish identification of the nonparametric joint distributions for each of the permanent state and transitory state variables over time. The constructive identification allows for closed-form sample counterpart estimators.

We apply the identification and estimation method to the earnings dynamics of U.S. men using the Panel Survey of Income Dynamics (PSID). Our results show that the marginal distributions of log earnings of U.S. men are nonnormal, with significant skewness and fatter tails of both the permanent and transitory components of earnings than the normal. We also find earnings dynamics very different than the normal, for our results show that the likelihood of remaining in a lower tail of the permanent earnings distribution does not fall over time as much, suggesting considerably more earnings mobility than would be found with a multivariate normality assumption. Another important finding from our empirical analysis is that the estimates of the marginal distributions as well of persistence and dynamics of permanent earnings are very sensitive to the degree of persistence in the transitory component. We find evidence for the existence of higher-order ARMA processes in the transitory component and that, with such higher-order processes, the permanent component of earnings has much less variability in marginal distributions and less mobility over time. Thus the transitory component makes a much stronger relative contribution to the marginal earnings distributions and to earnings mobility than in much of the prior literature, which often allows much less persistence in the transitory component. We also consider earnings dynamics in subsamples of men with strong labor force attachment and of married men, finding both subsamples to have lower variances of permanent and transitory shocks than for the full population but also more earnings mobility than that population.

As for future research, further generalizations of the state space model would be useful but the restrictions on the permanent state model and the transitory state model cannot be relaxed to the full extent simultaneously, because the permanent and transitory states are unobserved and cannot be distinguished without model restrictions. But there are a couple of directions for future research. One direction is to partially relax the independence of the innovation η_t in the permanent state transition, for instance, by accommodating heteroskedastic or dependently skewed distributions of permanent shocks, while keeping the Martingale feature of the canonical model. The other direction is to relax the semiparametric specification of the ARMA model for the transitory state variable by accommodating general nonparametric ARMA processes. Both directions are desirable, but it is essential to maintain nonparametric distributional assumptions and higher orders of the ARMA specification for the transitory state process as we stressed in discussing the empirical results.

As for applications, it would be helpful to modify our model to allow for changes in the earnings dynamic process with calendar time, for the growing literature on whether earnings volatility has been growing in the U.S. over time has only used simpler models of that process. An extension of our model to the earnings dynamics of women would also be of interest, for that would require adding a process for moving in and out of a zero-earnings state. Finally, applying our model to data sets drawn from administrative records (Social Security earnings, Unemployment Insurance earnings) would, given the large sample sizes of those data sets, allow more precise estimates of the distributions of the components, particularly in the tails.

Supplementary Material

The supplementary material contains large sample theories of the proposed estimator (Appendix C) and additional results of the application (Appendix D).

References

Abowd, J. and D. Card (1989) On the Covariance Structure of Earnings and Hours Changes. *Econometrica*, 57 (2), pp. 411–445.

Arellano, M., R. Blundell, and S. Bonhomme (2017) Earnings and Consumption Dynamics: A Nonlinear Panel Data Framework. *Econometrica*, 85 (3), pp. 693–734.

Blundell, R., L. Pistaferri, and I. Preston (2008) Consumption Inequality and Partial Insurance. *American Economic Review*, 98 (5), pp. 1887–1921.

Bonhomme, S. and J.-M. Robin (2010) Generalized Non-Parametric Deconvolution with an Application to Earnings Dynamics. *Review of Economic Studies*, 77 (2), pp. 491–533.

Botosaru, I. and Y. Sasaki (2018) Nonparametric Heteroskedasticity in Persistent Panel Processes: An Application to Earnings Dynamics. *Journal of Econometrics*, 203 (2), pp.283–296.

Browning, M., M. Ejrnaes, and J. Alvarez (2010) Modeling Income Processes with Lots of Heterogeneity. *Review of Economic Studies*, 77 (3), pp. 1353–81.

Čížek, P., W.K. Härdle, and R. Weron (2011) Statistical Tools for Finance and Insurance. Springer.

Delaigle, A. and I. Gijbels (2004) Practical Bandwidth Selection in Deconvolution Kernel Density Estimation. *Computational Statistics & Data Analysis*, 45 (2), pp. 249–267.

D'Haultfoeuille, X. (2011) On the Completeness Condition in Nonparametric Instrumental Problems. *Econometric Theory*, 27 (3), pp. 460–471.

Evdokimov, K. (2010) Identication and Estimation of a Nonparametric Panel Data Model with Unobserved Heterogeneity. Working Paper.

Evdokimov, K. and H. White (2012) Some Extensions of a Lemma of Kotlarski. *Econometric Theory*, 28 (4), pp. 925–932.

Horowitz, J.L. and M. Markatou (1996) Semiparametric Estimation of Regression Models for Panel Data. *Review of Economic Studies*, 63 (1), pp. 145–168.

Hryshko, D. (2012) Labor Income Profiles Are Not Heterogeneous: Evidence from Income Growth Rates. *Quantitative Economics*, 3 (3), pp. 177–209.

Hu Y. and Y. Sasaki (2017) Identification of Paired Nonseparable Measurement Error Models. *Econometric Theory,* 33 (4), pp. 955–979.

Hu, Y. and S.M. Schennach (2008) Instrumental Variable Treatment of Nonclassical Measurement Error Models. *Econometrica,* 76 (1), pp. 195–216.

Hu, Y., S.M. Schennach, and J.-L. Shiu (2018) Identification of Nonparametric Regression Models with Nonclassical Measurement Errors. Working Paper.

Kotlarski, I. (1967) On Characterizing The Gamma And The Normal Distribution. *Pacific Journal of Mathematics*, 20 (1), 69–76.

Kato, K., Y. Sasaki, and T. Ura (2018) Inference based on Kotlarski's Identity. arXiv:1808.09375

Li, T. and Q. Vuong (1998) "Nonparametric Estimation of the Measurement Error Model Using Multiple Indicators," *Journal of Multivariate Analysis*, 65 (2), 139–165.

Lillard, L. and R. Willis (1978) Dynamic Aspects of Earnings Mobility. *Econometrica*, 46 (3), pp. 985–1012.

MaCurdy, T. (1982) The Use of Time Series Processes to Model the Error Structure of Earnings in a Longitudinal Data Analysis. *Journal of Econometrics*, 18 (1), pp. 83–114.

Meghir, C. and L. Pistaferri (2004) Income Variance Dynamics and Heterogeneity. *Econometrica*, 72 (1), pp. 1–32.

Moffitt, R. and P. Gottschalk (2014) Trends in the Transitory Variance of Male Earnings: Methods and Evidence. *Journal of Human Resources*, 47 (4), pp. 204–36.

Moffitt, R. and S. Zhang (2018) Income Volatility and the PSID: Past Research and New Results. *American Economic Association Papers and Proceedings*, 108 (5) 277–280.

Newey, W.K. and D. McFadden (1994) Large Sample Estimation and Hypothesis Testing. *Handbook of Econometrics,* 4, Ch. 36, pp. 2111–2245.

Prakasa-Rao, B. L. S. (1992). Identifiability in Stochastic Models: Characterization of Probability Distributions. Probability and Mathematical Statistics. Academic Press, Boston.

Reiersol, O. (1950) Identifiability of a Linear Relation Between Variables Which are Subject to Error. *Econometrica*, 18 (4), pp. 375–389.

Schennach, S.M. (2004) Estimation of Nonlinear Models with Measurement Error. *Econometrica*, 72 (1), 33–75.

Schennach, S.M. and Y. Hu (2013) Nonparametric Identification and Semiparametric Estimation of Classical Measurement Error Models Without Side Information. *Journal of the American Statistical Association*, 108 (501), pp. 177–186.

Sample Definition	Age	Wage Salary	Wage Rate	Weeks Worked	Marital Status	Number of Observations
Baseline	30	34,090	12.706	46.583	0.751	1,320
		(19,711)	(4,879)	(8.313)	(0.433)	
	40	46,016	14.296	46.732	0.838	1,185
		(43,701)	(6.198)	(7.596)	(0.369)	
	50	46,923	14.682	45.680	0.887	984
		(37,883)	(5.872)	(8.850)	(0.317)	
Strong Labor	30	35,709	12.927	48.950	0.766	1,183
Force Attachment		(19,535)	(5.010)	(2.376)	(0.424)	
	40	48,123	14.490	48.715	0.849	1,076
		(44,950)	(6.270)	(2.336)	(0.359)	
	50	49,053	14.702	48.384	0.890	864
		(38,815)	(5.826)	(2.490)	(0.313)	
Married	30	35,092	13.250	47.018	1.000	991
		(19,145)	(4.959)	(7.499)	(0.000)	
	40	47,619	14.525	47.050	1.000	993
		(45,844)	(6.217)	(7.084)	(0.000)	
	50	47,929	14.898	45.647	1.000	873
		(38,674)	(5.778)	(8.923)	(0.000)	

Table 1: Summary statistics of wage salary, wage rate, weeks worked, and marital status. The numbers indicate the sample averages and the numbers in parentheses indicate the sample standard deviations. The statistics are provided under each of the three age groups, 30, 40, and 50, and uncer each of the three sample definitions: the baseline sample of workers, the sample of workers with strong labor force attachment (defined as 40 weeks or more of work in the previous year), and the sample of married workers. The currency units are real 1996 US dollars, deflated by the CPI-U-RS price deflator.

Baseline					
ARMA(0,0)	Mean	SD	Skewness	Kurtosis	H_0 : Kurtosis$\leqslant$3
U_{30}	0.000	0.441	-1.105	7.665	p-value = 0.002
	(0.016)	(0.018)	(0.284)	(1.644)	
U_{40}	-0.000	0.512	-1.136	9.748	p-value = 0.002
	(0.020)	(0.028)	(0.437)	(2.335)	
U_{50}	0.000	0.547	-1.341	11.169	p-value = 0.000
	(0.024)	(0.034)	(0.469)	(2.211)	
V_{30}	-0.000	0.297	-4.664	39.881	p-value = 0.000
	(0.012)	(0.029)	(0.721)	(8.407)	
V_{40}	0.000	0.227	-2.012	37.239	p-value = 0.499
	(0.013)	(0.040)	(74.459)	(13340.516)	
V_{50}	-0.000	0.222	-6.195	49.806	p-value = 0.487
	(1.422)	(3.695)	(1.303)	(1407.858)	

Table 2: Estimated distributional indices under the ARMA(0,0) model for the baseline sample. The indices include the mean, the standard deviation, the skewness, and the kurtosis. The numbers in parentheses indicate the standard errors of the respective estimates. The last column shows the p-value of the one-sided test of the null hypothesis that kurtosis is less than equal to three, against the alternative hypothesis that it is greater than three.

Baseline: Time-Varying Coefficients

ARMA(1,1)	Mean	SD	Skewness	Kurtosis	H_0 : Kurtosis$\leqslant$3
U_{30}	-0.000	0.391	-0.750	6.721	p-value = 0.068
	(0.016)	(0.016)	(0.367)	(2.494)	
U_{40}	-0.000	0.458	-0.506	6.858	p-value = 0.000
	(0.019)	(0.021)	(0.333)	(1.172)	
U_{50}	0.000	0.492	-0.928	9.681	p-value = 0.011
	(0.025)	(0.031)	(0.523)	(2.900)	
V_{30}	0.000	0.328	-3.991	28.236	p-value = 0.000
	(0.014)	(0.029)	(0.499)	(4.681)	
V_{40}	0.000	0.258	-3.016	27.830	p-value = 0.000
	(0.014)	(0.027)	(1.081)	(7.260)	
V_{50}	-0.000	0.232	-5.239	48.314	p-value = 0.362
	(0.018)	(0.039)	(4.439)	(128.250)	

Baseline: Time-Constant Coefficients

ARMA(1,1)	Mean	SD	Skewness	Kurtosis	H_0 : Kurtosis$\leqslant$3
U_{30}	-0.000	0.388	-0.718	6.672	p-value = 0.076
	(0.017)	(0.016)	(0.372)	(2.560)	
U_{40}	-0.000	0.458	-0.509	6.886	p-value = 0.000
	(0.019)	(0.021)	(0.339)	(1.170)	
U_{50}	0.000	0.493	-0.914	9.682	p-value = 0.009
	(0.025)	(0.030)	(0.522)	(2.822)	
V_{30}	0.000	0.332	-3.947	27.514	p-value = 0.000
	(0.015)	(0.028)	(0.491)	(4.670)	
V_{40}	0.000	0.258	-2.996	27.508	p-value = 0.000
	(0.014)	(0.027)	(1.075)	(7.096)	
V_{50}	-0.000	0.231	-5.411	48.554	p-value = 0.499
	(0.017)	(0.039)	(51.590)	(16379.615)	

Table 3: Estimated distributional indices under the ARMA(1,1) model for the baseline sample. The indices include the mean, the standard deviation, the skewness, and the kurtosis. The numbers in parentheses indicate the standard errors of the respective estimates. The last column shows the p-value of the one-sided test of the null hypothesis that kurtosis is less than equal to three, against the alternative hypothesis that it is greater than three.

Baseline: Time-Varying Coefficients

ARMA(2,2)	Mean	SD	Skewness	Kurtosis	H_0 : Kurtosis$\leqslant$3
U_{30}	0.000	0.394	-0.829	6.737	p-value = 0.040
	(0.018)	(0.018)	(0.352)	(2.136)	
U_{40}	0.000	0.433	-0.331	8.027	p-value = 0.000
	(0.020)	(0.022)	(0.410)	(1.403)	
U_{50}	0.000	0.457	-0.841	12.093	p-value = 0.009
	(0.031)	(0.034)	(0.708)	(3.868)	
V_{30}	-0.000	0.321	-4.246	32.980	p-value = 0.000
	(0.016)	(0.032)	(0.638)	(6.267)	
V_{40}	-0.000	0.284	-3.371	23.843	p-value = 0.000
	(0.016)	(0.027)	(0.929)	(5.346)	
V_{50}	-0.000	0.270	-4.619	35.106	p-value = 0.053
	(0.027)	(0.039)	(2.309)	(19.825)	

Baseline: Time-Invariant Coefficients

ARMA(2,2)	Mean	SD	Skewness	Kurtosis	H_0 : Kurtosis$\leqslant$3
U_{30}	0.000	0.391	-0.806	6.628	p-value = 0.049
	(0.019)	(0.019)	(0.370)	(2.196)	
U_{40}	0.000	0.432	-0.324	8.074	p-value = 0.000
	(0.021)	(0.023)	(0.416)	(1.431)	
U_{50}	0.000	0.464	-0.854	11.766	p-value = 0.008
	(0.028)	(0.033)	(0.686)	(3.670)	
V_{30}	-0.000	0.325	-4.170	32.056	p-value = 0.000
	(0.017)	(0.033)	(0.657)	(6.435)	
V_{40}	-0.000	0.285	-3.365	23.471	p-value = 0.000
	(0.017)	(0.028)	(0.945)	(5.238)	
V_{50}	-0.000	0.258	-4.986	39.363	p-value = 0.347
	(0.022)	(0.037)	(2.164)	(92.165)	

Table 4: Estimated distributional indices under the ARMA(2,2) model for the baseline sample. The indices include the mean, the standard deviation, the skewness, and the kurtosis. The numbers in parentheses indicate the standard errors of the respective estimates. The last column shows the p-value of the one-sided test of the null hypothesis that kurtosis is less than equal to three, against the alternative hypothesis that it is greater than three.

Baseline: Time-Varying Coefficients

ARMA(3,3)	Mean	SD	Skewness	Kurtosis	H_0 : Kurtosis$\leqslant$3
U_{30}	0.000	0.378	-0.949	7.307	p-value = 0.124
	(0.020)	(0.021)	(0.497)	(3.735)	
U_{40}	-0.000	0.418	-0.293	8.846	p-value = 0.000
	(0.024)	(0.023)	(0.453)	(1.465)	
U_{50}	0.000	0.455	-0.178	6.954	p-value = 0.002
	(0.031)	(0.029)	(0.432)	(1.403)	
V_{30}	-0.000	0.314	-4.260	31.880	p-value = 0.000
	(0.018)	(0.035)	(0.696)	(6.801)	
V_{40}	-0.000	0.283	-3.290	24.914	p-value = 0.000
	(0.021)	(0.028)	(0.879)	(6.106)	
V_{50}	-0.000	0.246	-5.946	47.363	p-value = 0.500
	(0.026)	(0.049)	(3910.480)	(833223.987)	

Baseline: Time-Invariant Coefficients

ARMA(3,3)	Mean	SD	Skewness	Kurtosis	H_0 : Kurtosis$\leqslant$3
U_{30}	0.000	0.376	-0.953	7.371	p-value = 0.134
	(0.021)	(0.022)	(0.526)	(3.950)	
U_{40}	-0.000	0.417	-0.289	8.876	p-value = 0.000
	(0.024)	(0.023)	(0.463)	(1.463)	
U_{50}	0.000	0.457	-0.198	6.952	p-value = 0.003
	(0.029)	(0.028)	(0.433)	(1.434)	
V_{30}	-0.000	0.317	-4.169	30.937	p-value = 0.000
	(0.020)	(0.035)	(0.711)	(6.917)	
V_{40}	-0.000	0.284	-3.275	24.660	p-value = 0.000
	(0.021)	(0.028)	(0.871)	(6.043)	
V_{50}	-0.000	0.242	-6.101	49.495	p-value = 0.496
	(0.023)	(0.046)	(84.350)	(4392.234)	

Table 5: Estimated distributional indices under the ARMA(3,3) model. The indices include the mean, the standard deviation, the skewness, and the kurtosis. The numbers in parentheses indicate the standard errors of the respective estimates. The last column shows the p-value of the one-sided test of the null hypothesis that kurtosis is less than equal to three, against the alternative hypothesis that it is greater than three.

Baseline: Time-Varying Coefficients

ARMA(4,4)	Mean	SD	Skewness	Kurtosis	H_0 : Kurtosis$\leqslant$3
U_{30}	-0.000	0.360	-0.689	6.071	p-value = 0.197
	(0.020)	(0.022)	(0.496)	(3.603)	
U_{40}	-0.000	0.420	-0.311	7.804	p-value = 0.001
	(0.024)	(0.024)	(0.453)	(1.520)	
U_{50}	0.000	0.429	-0.612	7.068	p-value = 0.011
	(0.032)	(0.030)	(0.449)	(1.769)	
V_{30}	0.000	0.318	-4.434	34.961	p-value = 0.000
	(0.020)	(0.039)	(0.766)	(9.531)	
V_{40}	-0.000	0.288	-3.431	31.704	p-value = 0.001
	(0.020)	(0.035)	(1.650)	(9.283)	
V_{50}	-0.000	0.287	-3.711	31.328	p-value = 0.132
	(0.029)	(0.045)	(2.103)	(25.319)	

Baseline: Time-Invariant Coefficients

ARMA(4,4)	Mean	SD	Skewness	Kurtosis	H_0 : Kurtosis$\leqslant$3
U_{30}	-0.000	0.355	-0.640	5.686	p-value = 0.237
	(0.022)	(0.023)	(0.519)	(3.748)	
U_{40}	-0.000	0.419	-0.305	7.764	p-value = 0.001
	(0.025)	(0.024)	(0.451)	(1.539)	
U_{50}	0.000	0.436	-0.581	6.992	p-value = 0.011
	(0.029)	(0.028)	(0.434)	(1.734)	
V_{30}	0.000	0.323	-4.325	33.745	p-value = 0.000
	(0.022)	(0.040)	(0.817)	(9.336)	
V_{40}	-0.000	0.289	-3.418	31.635	p-value = 0.002
	(0.021)	(0.036)	(1.770)	(9.894)	
V_{50}	-0.000	0.276	-4.163	34.864	p-value = 0.456
	(0.024)	(0.044)	(8.261)	(289.995)	

Table 6: Estimated distributional indices under the ARMA(4,4) model. The indices include the mean, the standard deviation, the skewness, and the kurtosis. The numbers in parentheses indicate the standard errors of the respective estimates. The last column shows the p-value of the one-sided test of the null hypothesis that kurtosis is less than equal to three, against the alternative hypothesis that it is greater than three.

Strong Labor Force Attachment

ARMA(0,0)	Mean	SD	Skewness	Kurtosis	H_0 : Kurtosis$\leqslant$3
U_{30}	-0.000	0.332	-0.586	2.992	p-value = 0.513
	(0.013)	(0.009)	(0.096)	(0.249)	
U_{40}	0.000	0.379	-0.399	3.240	p-value = 0.112
	(0.015)	(0.011)	(0.093)	(0.197)	
U_{50}	-0.000	0.382	-0.232	3.212	p-value = 0.326
	(0.019)	(0.014)	(0.151)	(0.472)	
V_{30}	0.000	0.187	-1.436	9.240	p-value = 0.001
	(0.009)	(0.012)	(0.439)	(2.073)	
V_{40}	-0.000	0.152	-2.525	24.713	p-value = 0.005
	(0.008)	(0.016)	(1.108)	(8.389)	
V_{50}	-0.000	0.173	-4.332	39.945	p-value = 0.044
	(0.011)	(0.028)	(1.832)	(21.681)	

Table 7: Estimated distributional indices under the ARMA(0,0) model for the sub-sample of individuals with strong labor force attachment. The indices include the mean, the standard deviation, the skewness, and the kurtosis. The numbers in parentheses indicate the standard errors of the respective estimates. The last column shows the p-value of the one-sided test of the null hypothesis that kurtosis is less than equal to three, against the alternative hypothesis that it is greater than three.

Strong Labor Force Attachment: Time-Varying Coefficients

ARMA(1,1)	Mean	SD	Skewness	Kurtosis	H_0 : Kurtosis$\leqslant$3
U_{30}	0.000	0.316	-0.532	2.774	p-value = 0.790
	(0.014)	(0.010)	(0.107)	(0.281)	
U_{40}	-0.000	0.358	-0.556	3.401	p-value = 0.108
	(0.016)	(0.013)	(0.139)	(0.324)	
U_{50}	-0.000	0.371	-0.351	3.115	p-value = 0.413
	(0.023)	(0.017)	(0.158)	(0.527)	
V_{30}	-0.000	0.192	-1.323	7.837	p-value = 0.006
	(0.011)	(0.014)	(0.420)	(1.945)	
V_{40}	0.000	0.184	-1.031	11.349	p-value = 0.004
	(0.011)	(0.016)	(0.859)	(3.153)	
V_{50}	0.000	0.158	-4.770	64.914	p-value = 0.497
	(0.016)	(0.040)	(132.359)	(9677.548)	

Strong Labor Force Attachment: Time-Constant Coefficients

ARMA(1,1)	Mean	SD	Skewness	Kurtosis	H_0 : Kurtosis$\leqslant$3
U_{30}	0.000	0.315	-0.535	2.764	p-value = 0.794
	(0.015)	(0.010)	(0.109)	(0.288)	
U_{40}	-0.000	0.358	-0.560	3.405	p-value = 0.111
	(0.017)	(0.013)	(0.144)	(0.332)	
U_{50}	-0.000	0.370	-0.335	3.085	p-value = 0.435
	(0.0209)	(0.016)	(0.153)	(0.524)	
V_{30}	-0.000	0.193	-1.299	7.747	p-value = 0.010
	(0.011)	(0.014)	(0.436)	(2.035)	
V_{40}	0.000	0.185	-1.006	11.157	p-value = 0.004
	(0.011)	(0.017)	(0.863)	(3.122)	
V_{50}	0.000	0.160	-4.784	61.908	p-value = 0.484
	(0.014)	(0.037)	(39.663)	(1470.245)	

Table 8: Estimated distributional indices under the ARMA(1,1) model for the sub-sample of individuals with strong labor force attachment. The indices include the mean, the standard deviation, the skewness, and the kurtosis. The numbers in parentheses indicate the standard errors of the respective estimates. The last column shows the p-value of the one-sided test of the null hypothesis that kurtosis is less than equal to three, against the alternative hypothesis that it is greater than three.

Strong Labor Force Attachment: Time-Varying Coefficients

ARMA(2,2)	Mean	SD	Skewness	Kurtosis	H_0 : Kurtosis$\leqslant$3
U_{30}	0.000	0.305	-0.390	2.479	p-value = 0.993
	(0.015)	(0.010)	(0.105)	(0.210)	
U_{40}	0.000	0.331	-0.338	3.227	p-value = 0.269
	(0.018)	(0.013)	(0.146)	(0.369)	
U_{50}	0.000	0.334	-0.298	2.300	p-value = 0.661
	(0.024)	(0.020)	(0.304)	(1.688)	
V_{30}	-0.000	0.188	-1.503	7.165	p-value = 0.021
	(0.012)	(0.014)	(0.452)	(2.055)	
V_{40}	-0.000	0.197	-1.640	8.369	p-value = 0.016
	(0.014)	(0.017)	(0.522)	(2.518)	
V_{50}	-0.000	0.223	-2.306	25.355	p-value = 0.072
	(0.020)	(0.038)	(2.083)	(15.309)	

Strong Labor Force Attachment: Time-Invariant Coefficients

ARMA(2,2)	Mean	SD	Skewness	Kurtosis	H_0 : Kurtosis$\leqslant$3
U_{30}	0.000	0.303	-0.377	2.442	p-value = 0.995
	(0.016)	(0.010)	(0.110)	(0.219)	
U_{40}	0.000	0.330	-0.333	3.223	p-value = 0.281
	(0.018)	(0.013)	(0.147)	(0.385)	
U_{50}	0.000	0.341	-0.269	2.449	p-value = 0.660
	(0.022)	(0.018)	(0.253)	(1.331)	
V_{30}	-0.000	0.191	-1.513	7.077	p-value = 0.029
	(0.013)	(0.015)	(0.481)	(2.155)	
V_{40}	-0.000	0.199	-1.638	8.247	p-value = 0.017
	(0.014)	(0.017)	(0.526)	(2.469)	
V_{50}	0.000	0.212	-2.729	29.773	p-value = 0.049
	(0.017)	(0.037)	(2.113)	(16.189)	

Table 9: Estimated distributional indices under the ARMA(2,2) model for the sub-sample of individuals with strong labor force attachment. The indices include the mean, the standard deviation, the skewness, and the kurtosis. The numbers in parentheses indicate the standard errors of the respective estimates. The last column shows the p-value of the one-sided test of the null hypothesis that kurtosis is less than equal to three, against the alternative hypothesis that it is greater than three.

Strong Labor Force Attachment: Time-Varying Coefficients

ARMA(3,3)	Mean	SD	Skewness	Kurtosis	H_0 : Kurtosis$\leqslant$3
U_{30}	0.000	0.292	-0.451	2.459	p-value = 0.983
	(0.016)	(0.011)	(0.120)	(0.256)	
U_{40}	-0.000	0.330	-0.352	3.355	p-value = 0.182
	(0.019)	(0.013)	(0.148)	(0.390)	
U_{50}	0.000	0.345	-0.275	2.445	p-value = 0.768
	(0.023)	(0.017)	(0.195)	(0.758)	
V_{30}	-0.000	0.203	-1.202	5.341	p-value = 0.071
	(0.013)	(0.015)	(0.352)	(1.596)	
V_{40}	0.000	0.181	-1.675	10.175	p-value = 0.030
	(0.014)	(0.017)	(0.682)	(3.829)	
V_{50}	-0.000	0.218	-2.589	28.596	p-value = 0.037
	(0.017)	(0.037)	(1.976)	(14.329)	

Strong Labor Force Attachment: Time-Invariant Coefficients

ARMA(3,3)	Mean	SD	Skewness	Kurtosis	H_0 : Kurtosis$\leqslant$3
U_{30}	0.000	0.291	-0.451	2.456	p-value = 0.984
	(0.016)	(0.011)	(0.120)	(0.255)	
U_{40}	-0.000	0.330	-0.351	3.354	p-value = 0.183
	(0.019)	(0.013)	(0.148)	(0.391)	
U_{50}	0.000	0.345	-0.269	2.445	p-value = 0.768
	(0.023)	(0.017)	(0.192)	(0.757)	
V_{30}	-0.000	0.204	-1.196	5.305	p-value = 0.077
	(0.013)	(0.015)	(0.356)	(1.613)	
V_{40}	0.000	0.181	-1.676	10.167	p-value = 0.028
	(0.014)	(0.017)	(0.683)	(3.759)	
V_{50}	-0.000	0.218	-2.638	29.049	p-value = 0.032
	(0.017)	(0.038)	(1.965)	(14.085)	

Table 10: Estimated distributional indices under the ARMA(3,3) model for the sub-sample of individuals with strong labor force attachment. The indices include the mean, the standard deviation, the skewness, and the kurtosis. The numbers in parentheses indicate the standard errors of the respective estimates. The last column shows the p-value of the one-sided test of the null hypothesis that kurtosis is less than equal to three, against the alternative hypothesis that it is greater than three.

Strong Labor Force Attachment: Time-Varying Coefficients

ARMA(4,4)	Mean	SD	Skewness	Kurtosis	H_0 : Kurtosis$\leqslant$3
U_{30}	-0.000	0.277	-0.315	2.371	p-value = 0.978
	(0.017)	(0.012)	(0.139)	(0.313)	
U_{40}	0.000	0.323	-0.617	3.473	p-value = 0.180
	(0.020)	(0.015)	(0.169)	(0.517)	
U_{50}	0.000	0.337	-0.280	2.021	p-value = 0.895
	(0.025)	(0.018)	(0.184)	(0.779)	
V_{30}	0.000	0.205	-1.355	5.507	p-value = 0.080
	(0.014)	(0.016)	(0.368)	(1.781)	
V_{40}	0.000	0.191	-0.999	9.042	p-value = 0.044
	(0.015)	(0.019)	(0.710)	(3.537)	
V_{50}	-0.000	0.228	-2.337	28.495	p-value = 0.041
	(0.021)	(0.043)	(2.273)	(14.647)	

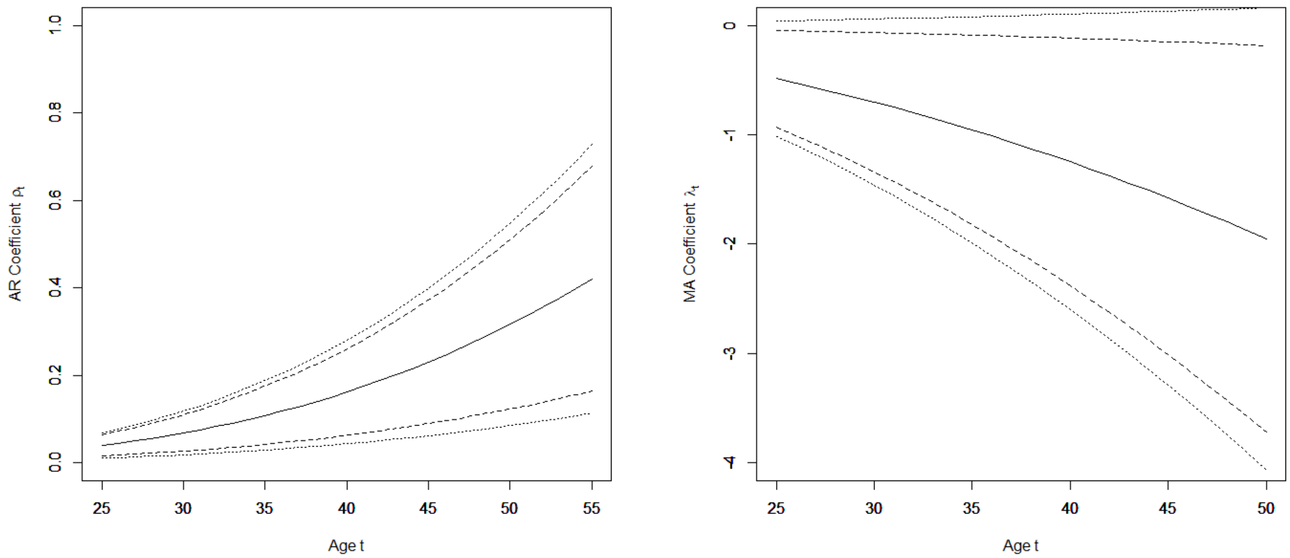
Strong Labor Force Attachment: Time-Invariant Coefficients

ARMA(4,4)	Mean	SD	Skewness	Kurtosis	H_0 : Kurtosis$\leqslant$3
U_{30}	-0.000	0.277	-0.309	2.359	p-value = 0.977
	(0.017)	(0.012)	(0.142)	(0.321)	
U_{40}	0.000	0.323	-0.618	3.473	p-value = 0.176
	(0.020)	(0.015)	(0.167)	(0.507)	
U_{50}	0.000	0.337	-0.270	2.021	p-value = 0.890
	(0.026)	(0.018)	(0.184)	(0.799)	
V_{30}	0.000	0.206	-1.357	5.483	p-value = 0.080
	(0.015)	(0.016)	(0.363)	(1.764)	
V_{40}	-0.000	0.191	-0.997	9.031	p-value = 0.043
	(0.015)	(0.019)	(0.714)	(3.520)	
V_{50}	-0.000	0.227	-2.386	28.827	p-value = 0.097
	(0.021)	(0.043)	(2.329)	(19.870)	

Table 11: Estimated distributional indices under the ARMA(4,4) model for the sub-sample of individuals with strong labor force attachment. The indices include the mean, the standard deviation, the skewness, and the kurtosis. The numbers in parentheses indicate the standard errors of the respective estimates. The last column shows the p-value of the one-sided test of the null hypothesis that kurtosis is less than equal to three, against the alternative hypothesis that it is greater than three.

Baseline Sample

AR Parameter $\widehat{\rho}_t$ MA Parameter $\widehat{\lambda}_t$



Sub-Sample of Workers with Strong Labor Force Attachment

AR Parameter $\widehat{\rho}_t$ MA Parameter $\widehat{\lambda}_t$

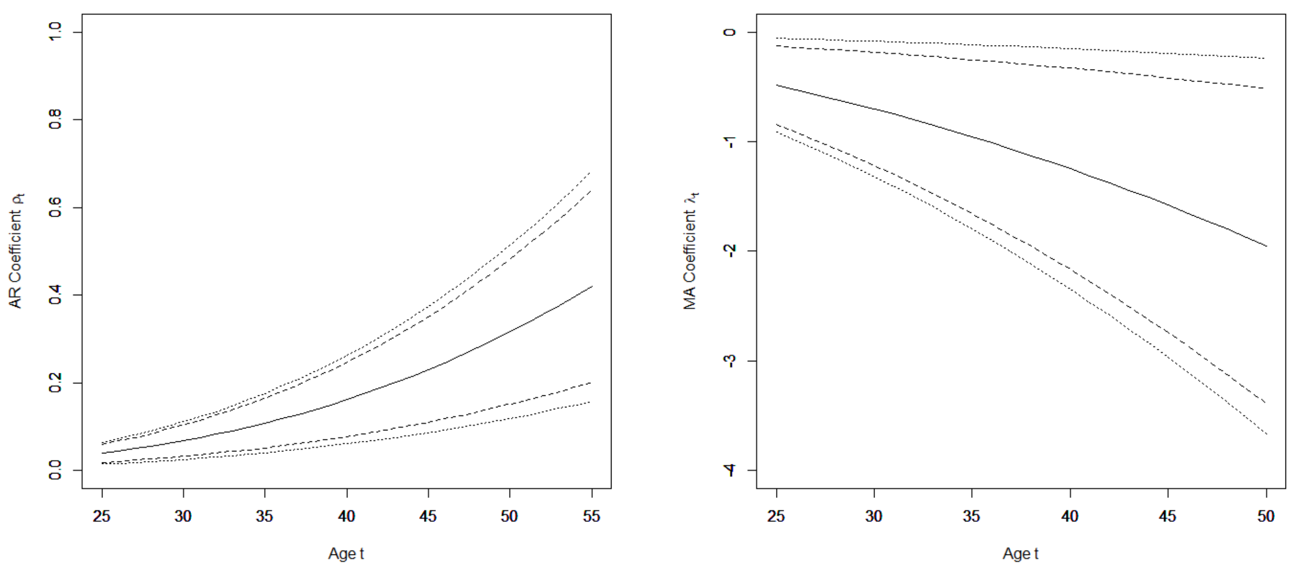


Figure 1: Estimates of the AR parameter (left) and the MA parameter (right) under the ARMA(1,1) specification for the baseline sample (top) and the sub-sample of workers with strong labor force attachment (bottom). The dashed and dotted curves indicate the boundary of 90 percent and 95 percent confidence intervals, respectively.

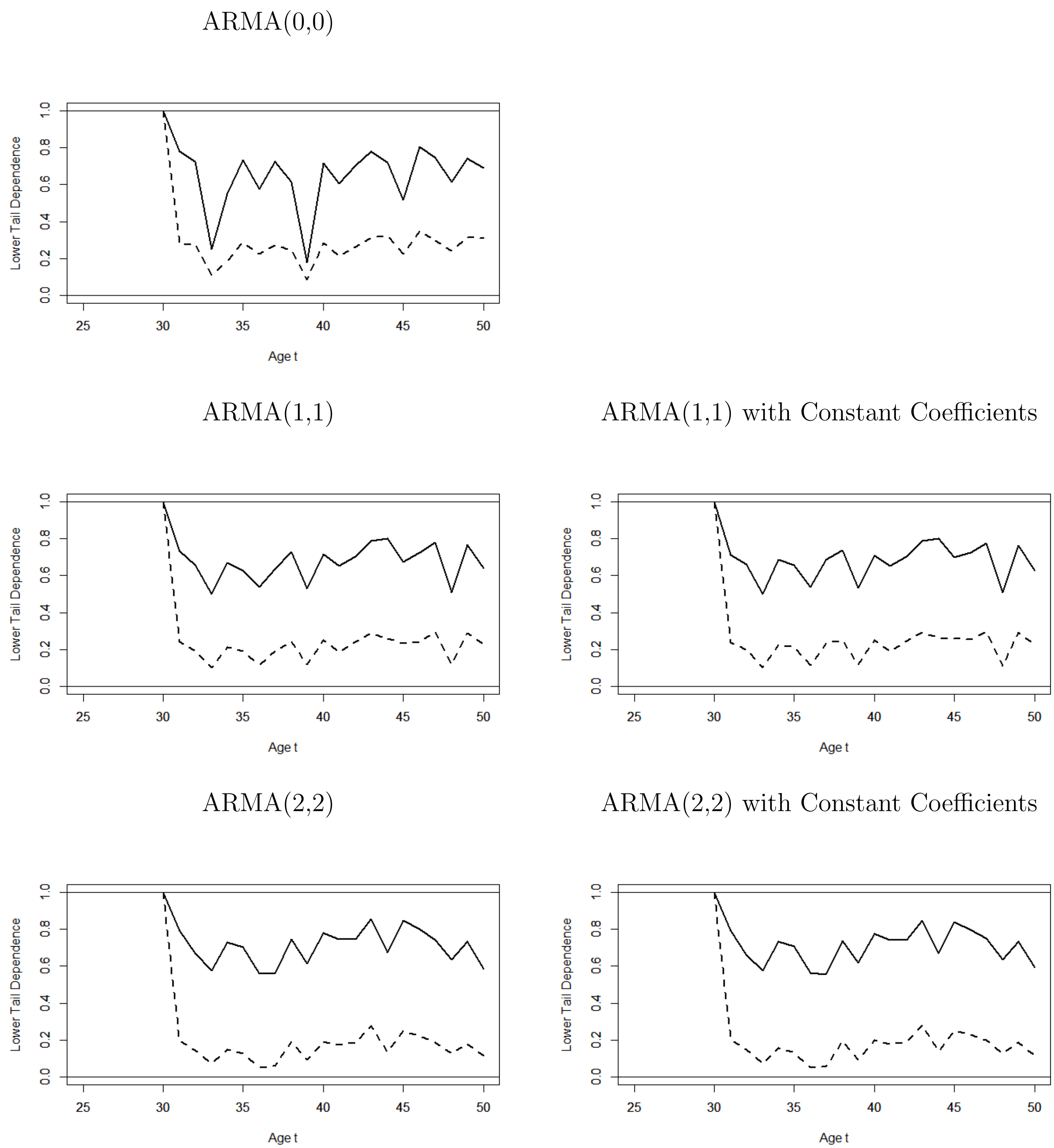


Figure 2a: Trajectories of the lower tail dependence measure $\lambda^{l}_{30,t}(0.01) = P(U_t \leq F^{-1}_{U_t}(0.01)|U_{30} \leq F^{-1}_{U_{30}}(0.01))$ of permanent earnings following the event of permanent earnings less than or equal to the 1 percentile at age 30. The results are based on the baseline sample. The solid lines represent the trajectories under our semiparametric model, while the dashed lines represent those under the bivariate normal distribution. The results are displayed under each of the ARMA(0,0), ARMA(1,1), and ARMA(2,2) specifications with time-varying coefficients and time-invariant coefficients.

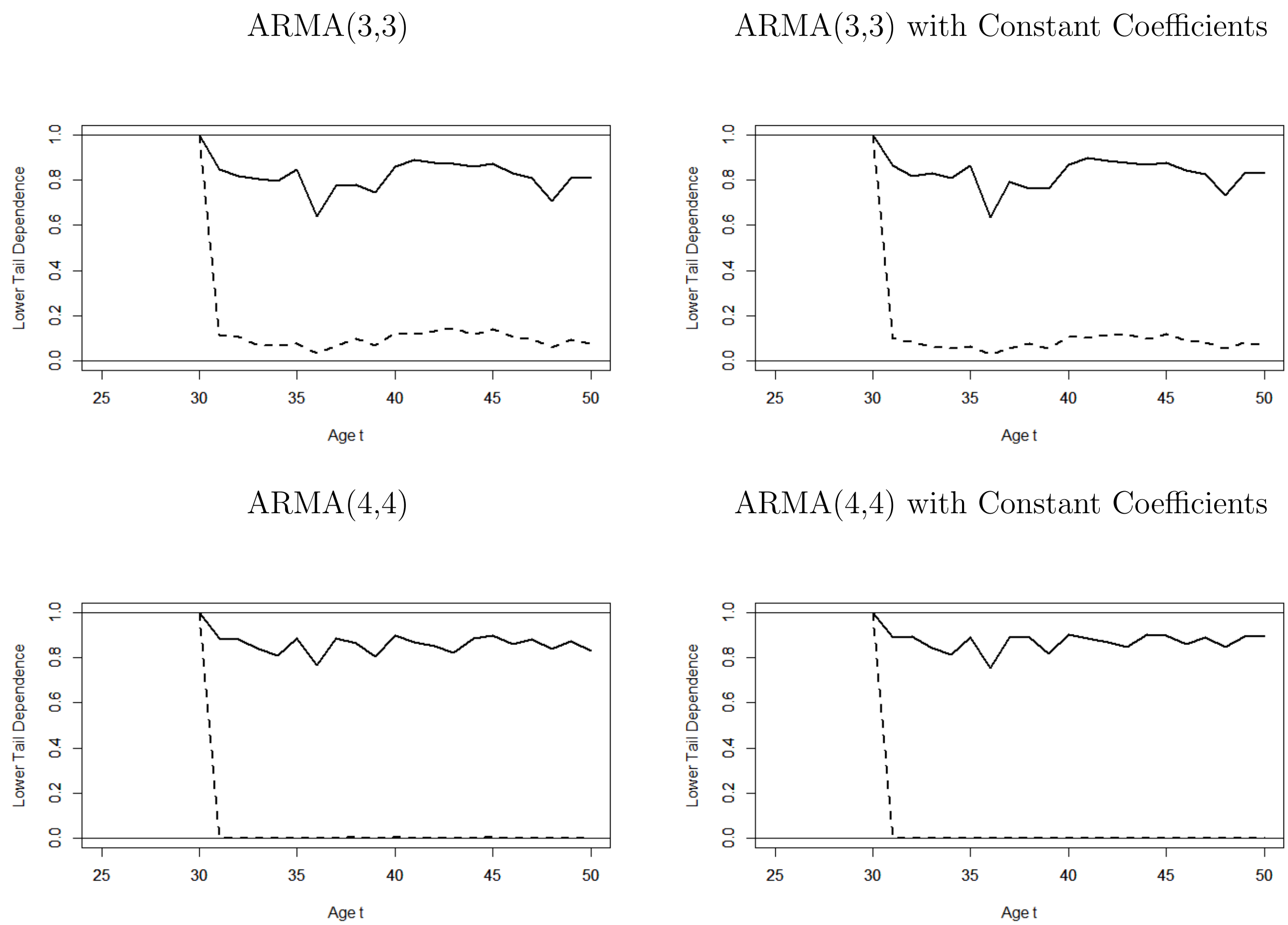


Figure 2b: Trajectories of the lower tail dependence measure $\lambda^{l}_{30,t}(0.01) = P(U_t \leq F^{-1}_{U_t}(0.01) | U_{30} \leq F^{-1}_{U_{30}}(0.01))$ of permanent earnings following the event of permanent earnings less than or equal to the 1 percentile at age 30. The results are based on the baseline sample. The solid lines represent the trajectories under our semiparametric model, while the dashed lines represent those under the bivariate normal distribution. The results are displayed under each of the ARMA(3,3) and ARMA(4,4) specifications with time-varying coefficients and time-invariant coefficients.

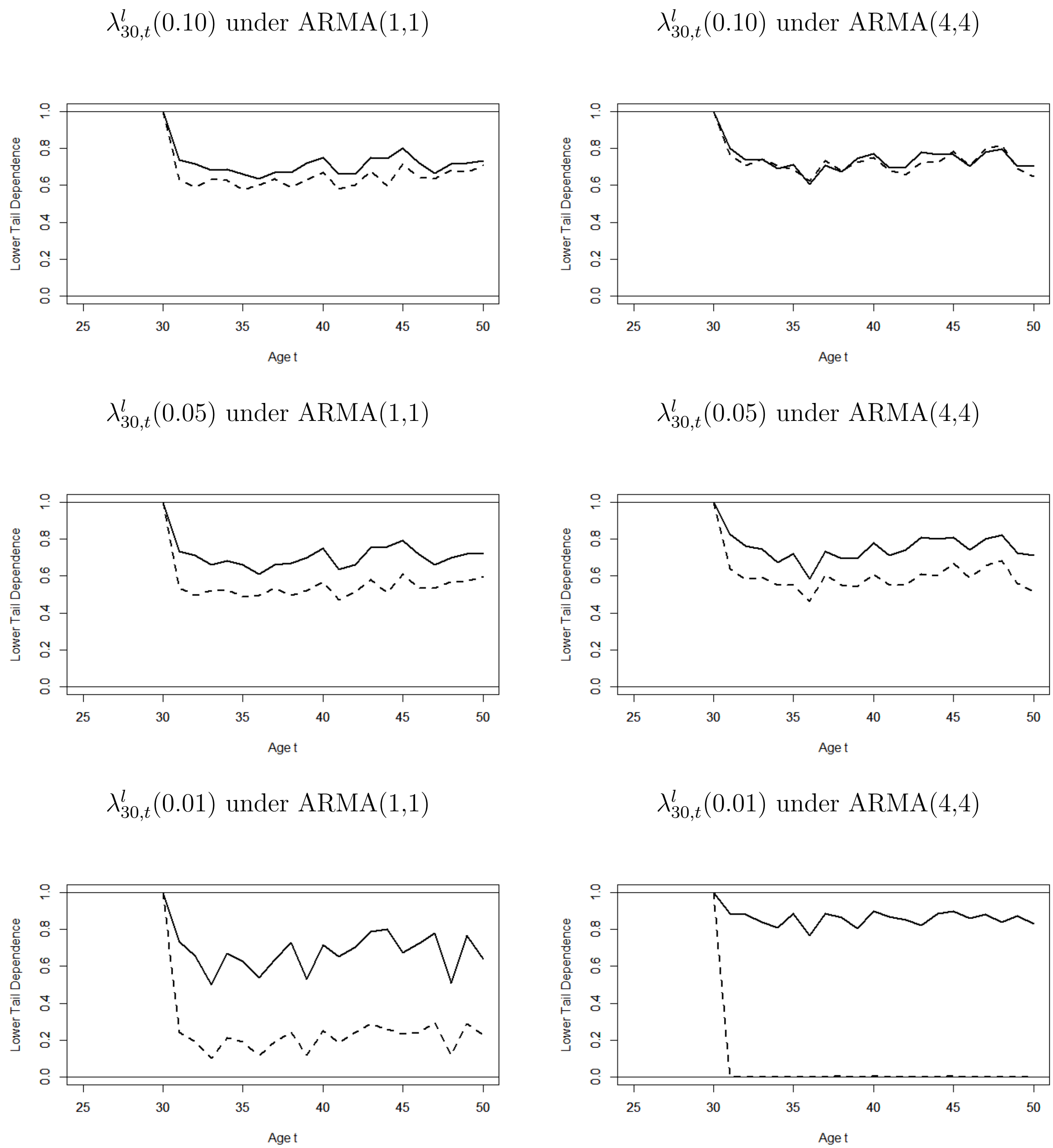


Figure 3: Trajectories of the lower tail dependence measure $\lambda^l_{30,t}(q) = P(U_t \leq F^{-1}_{U_t}(q) | U_{30} \leq F^{-1}_{U_{30}}(q))$ of permanent earnings following the event of permanent earnings less than or equal to the q-th quantile at age 30 for $q \in \{0.10, 0.05, 0.01\}$. The results are based on the baseline sample. The solid lines represent the trajectories under our semiparametric model, while the dashed lines represent those under the bivariate normal distribution. The results are displayed under each of the ARMA(1,1) and ARMA(4,4) specifications with time-varying coefficients.

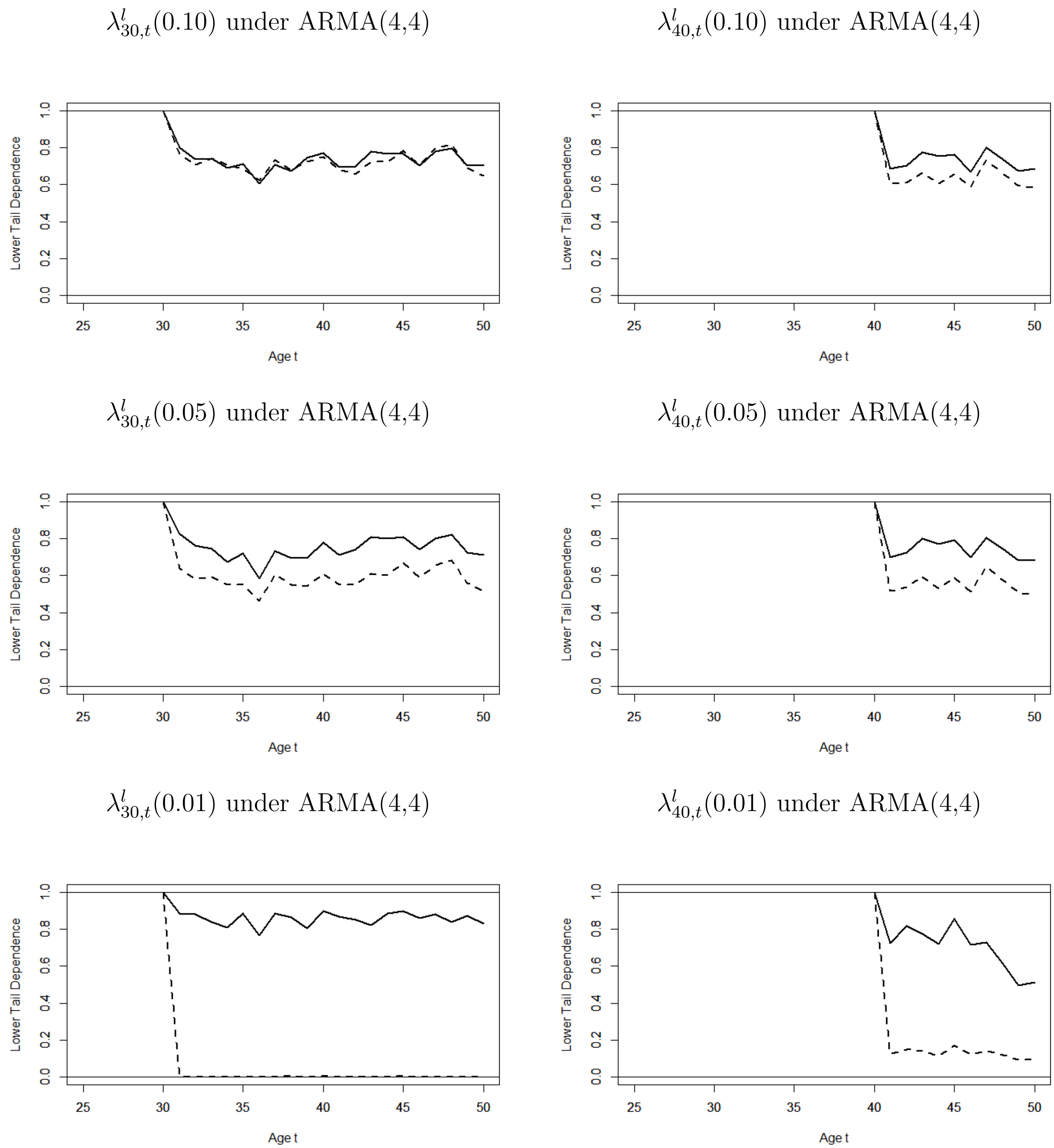


Figure 4: Trajectories of the lower tail dependence measures $\lambda^l_{30,t}(q) = P(U_t \leq F^{-1}_{U_t}(q)|U_{30} \leq F^{-1}_{U_{30}}(q))$ and $\lambda^l_{40,t}(q) = P(U_t \leq F^{-1}_{U_t}(q)|U_{40} \leq F^{-1}_{U_{40}}(q))$ of permanent earnings following the event of permanent earnings less than or equal to the q-th quantile at age 30 and 40, respectively, for $q \in \{0.10, 0.05, 0.01\}$. The results are based on the baseline sample. The solid lines represent the trajectories under our semiparametric model, while the dashed lines represent those under the bivariate normal distribution. The results are displayed under the ARMA(4,4) specification with time-varying coefficients.

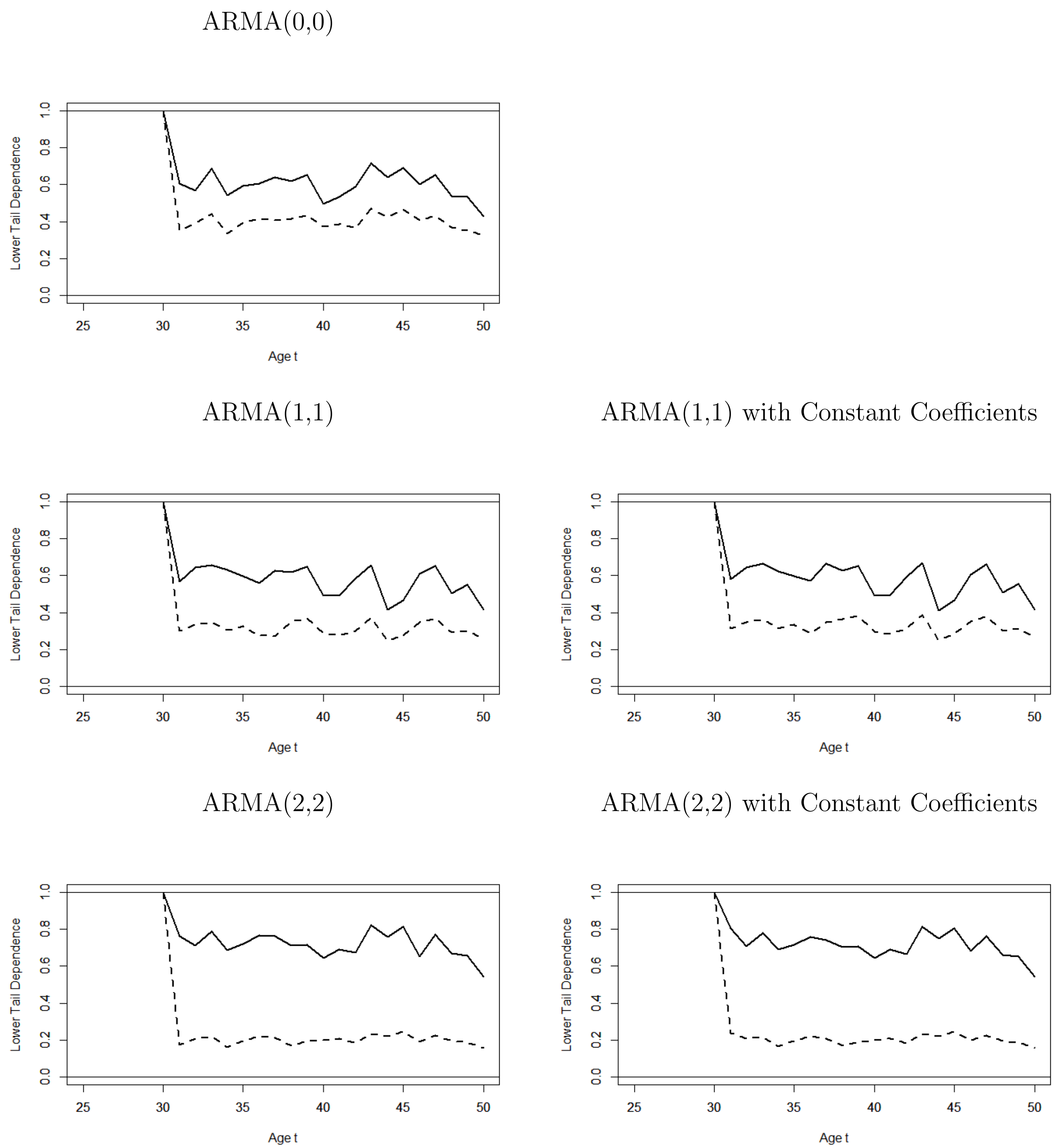


Figure 5a: Trajectories of the lower tail dependence measure $\lambda^l_{30,t}(0.01) = P(U_t \leq F^{-1}_{U_t}(0.01)|U_{30} \leq F^{-1}_{U_{30}}(0.01))$ of permanent earnings following the event of permanent earnings less than or equal to the 1 percentile at age 30. The sample consists of individuals with strong labor force attachment. The solid lines represent the trajectories under our semiparametric model, while the dashed lines represent those under the bivariate normal distribution. The results are displayed under each of the ARMA(0,0), ARMA(1,1), and ARMA(2,2) specifications with time-varying coefficients and time-invariant coefficients.

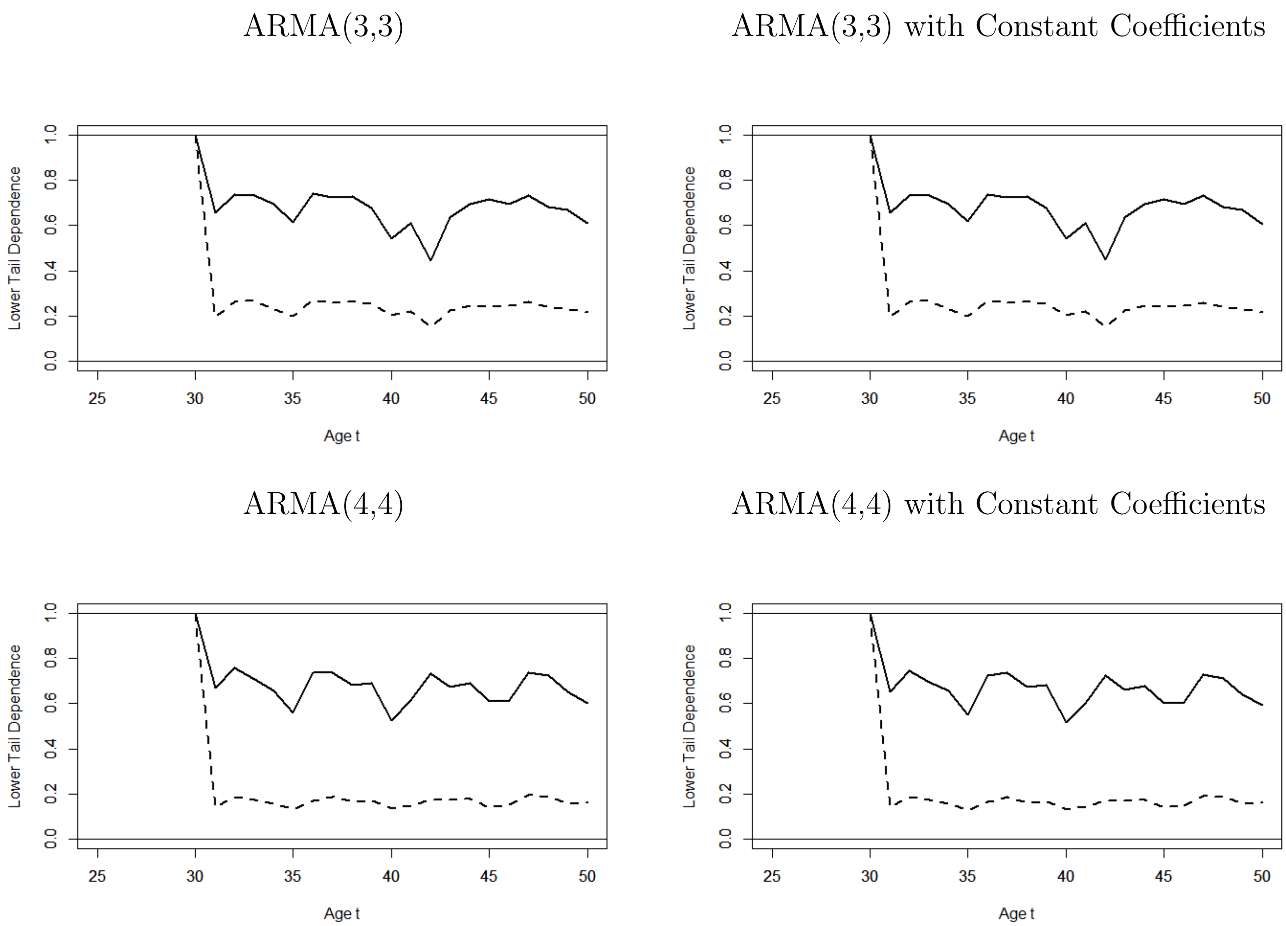


Figure 5b: Trajectories of the lower tail dependence measure $\lambda^l_{30,t}(0.01) = P(U_t \leq F^{-1}_{U_t}(0.01)|U_{30} \leq F^{-1}_{U_{30}}(0.01))$ of permanent earnings following the event of permanent earnings less than or equal to the 1 percentile at age 30. The sample consists of individuals with strong labor force attachment. The solid lines represent the trajectories under our semiparametric model, while the dashed lines represent those under the bivariate normal distribution. The results are displayed under each of the ARMA(3,3) and ARMA(4,4) specifications with time-varying coefficients and time-invariant coefficients.