

Identification of the Covariance Structure of Earnings using the GMM Estimator

Aedín Doris*, Donal O'Neill** & Olive Sweetman*

JEL Codes: J31, D31

Abstract

In this paper we study the performance of the GMM estimator in the context of the covariance structure of earnings. Using analytical and Monte Carlo techniques we examine the sensitivity of parameter identification to key features such as panel length, sample size, the degree of persistence of earnings shocks and the evolution of inequality over time. We show that the interaction of transitory persistence with the time pattern of inequality determines identification in these models and offer some practical recommendations that follow from our findings.

** Corresponding author; National University of Ireland Maynooth and IZA, Bonn. E-mail: donal.oneill@nuim.ie; tel.: 353-1-7083555; fax: 353-1-7083934; address: Rhetoric House, NUI Maynooth, Maynooth, Co. Kildare, Ireland.

*National University of Ireland Maynooth.

We would like to thank Peter Gottschalk for helpful comments on an earlier version of this paper, as well as seminar participants at NUI Maynooth, and participants at the *Income and Earnings Dynamics* conference at NUI Maynooth, the BHPS-2009 conference in Essex and the Irish Economics Association Annual Conference 2010, Belfast. We would also like to thank Steven Haider, Marco Lilla, Xavier Ramos, Denisa Sologon and Robert Valetta, who provided us with additional details of their work that were helpful when conducting our research. We gratefully acknowledge the financial support provided by the Irish Research Council for the Humanities and Social Sciences.

1. Introduction

In recent years there has been a rapid growth in the number of studies that have used the Generalised Method of Moments (GMM) estimator to estimate the covariance structure of earnings (e.g. Moffitt and Gottschalk (1995, 2002, 2008), Dickens (2000), Haider (2001), Ramos (2002), Baker and Solon (2003), Capellari (2004), Gustavsson (2004), Daly and Valetta (2007) and Kalwij and Alessie (2007)).¹ In these models, earnings are written as the sum of permanent and transitory components. The resulting parameter estimates are then used to construct measures of permanent and transitory inequality and to trace their evolution over time. Distinguishing between these two components is important because they have different policy implications; moreover the distinction can provide insight into the functioning of the labour market.

The GMM estimator uses panel data to estimate these models by matching the sample variances and covariances of earnings to their population counterparts. The model is identified from the long covariances. In these latter moments, the contribution of the transitory shock is negligible, which in turn allows researchers to recover the parameters associated with the permanent component. However long panels are not always available to researchers and as a result a number of recent studies, for example Ramos (2003), Doris et al. (2008), Cervini and Ramos (2008) and Sologon and O'Donoghue (2009), have been constrained to use relatively short panels, with eight or nine years of data. It is unclear whether panel lengths of this order are sufficient to identify these models. Although the performance of the GMM estimator has been evaluated elsewhere (e.g. Tauchen (1986), Hansen et al. (1996), Altonji and Segal (1996), Clark (1996), Stock and Wright (2000) and Pozzi (2003)),

¹ Dearden et al. (2006), Shin and Solon (2008) and Bonhomme and Robin (2008) also estimate the covariance structure of earnings, but do not use GMM.

as yet there has been no detailed study of the estimator for the type of earnings covariance models or the data structures often found in the empirical literature. Using both analytical and Monte Carlo techniques we consider identification of these models and discuss the consequences for estimation and inference.

Section 2 reviews the conditions needed for identification using the GMM estimator and Section 3 presents the earnings model considered in this paper. Section 4 examines the sensitivity of parameter identification to key features such as panel length, number of observations, the degree of persistence of earnings shocks and the trends in inequality. We discuss the circumstances under which traditional asymptotics provide a good approximation to the distribution of GMM estimators in these models. We show that identification depends crucially on the evolution of inequality over time and offer some practical recommendations that follow from our findings.

2. Identification Using GMM

The GMM approach to parameter estimation is now well established in the econometric literature, having been introduced by Hansen (1982). Hall (2005) provides a comprehensive discussion of the approach in a time-series context, while Cameron and Trivedi (2005) discuss the procedure in a microeconomic setting. GMM is based on the analogy principle whereby population moment conditions are replaced by their sample analogues. This in turn provides a system of equations which form the basis for the derivation of the Method of Moments estimator. Formally, suppose we have a $k \times 1$ vector m and a $p \times 1$ parameter vector θ such that for a given value θ_0 and data Y

$$E[m(Y; \theta_0)] = 0, \tag{1}$$

The GMM approach replaces the population expectation with the sample moments

$$\bar{m}(\theta) = \frac{1}{N} \sum_{i=1}^N m_i(y_i; \theta) \quad (2)$$

and chooses the value of θ that makes $\bar{m}(\theta)$ equal to or ‘close to’ zero. Formally the GMM estimator chooses the value of θ so as to minimise the criterion function

$$\bar{m}(\theta)' W_n \bar{m}(\theta), \quad (3)$$

where W_n is a positive semi-definite matrix that does not depend on θ . The GMM estimator will identify the model if the probability limit of the GMM criterion function is uniquely minimised at the true parameter vector, θ_0 . If a model is not identified, there exist at least two distinct data-generating processes (DGPs), characterised by different parameter vectors θ , which cannot be distinguished by any function of the data, even with infinitely large samples. It has been shown that GMM identifies θ_0 , provided that for all $\theta \neq \theta_0$, then $E[m(Y; \theta)]$ is not in the null-space of $W = \text{plim} W_n$ (Newey and McFadden, 1994). When W is positive-definite, this reduces to the more familiar moment condition:

$$E[m(Y; \theta_0)] = 0 \text{ and } E[m(Y; \theta)] \neq 0 \quad \forall \theta \neq \theta_0$$

A necessary condition for identification by the GMM estimator is that the number of moment conditions is at least as great as the number of parameters (the order condition). In addition, the rank condition requires that the information provided by the moment conditions must differ; that is, as the p components of θ vary in the neighbourhood of θ_0 , the k components of $m(Y; \theta)$ vary in p independent directions (see for example Hall, 2005, Chapter 3). We will use this condition to examine the identification of our models later in the paper.

Altonji and Segal (1996) and Clark (1996) consider the appropriate weighting matrix in relatively straightforward covariance models, and show that the matrix that is optimal in large samples can lead to biased results in small samples and that the use of the identity matrix is preferable. This approach has therefore become common practice when estimating the covariance structure of earnings. In order to ensure that our results are comparable to the applied literature, we also use the identity matrix in our analysis. Under suitable regularity and identification conditions, it can then be shown that the limiting distribution of our estimator $\hat{\theta}_{GMM}$ is as follows:

$$\sqrt{N} \left(\hat{\theta}_{GMM} - \theta_0 \right) \xrightarrow{d} N \left\{ 0, (G_0' G_0)^{-1} (G_0' S_0 G_0) (G_0' G_0)^{-1} \right\} \quad (4)$$

$$\text{where } G_0 = \text{plim} \frac{1}{N} \sum_{i=1}^N \left[\frac{\delta m_i}{\delta \theta} \Big|_{\theta_0} \right] \text{ and } S_0 = \text{plim} \frac{1}{N} \sum_{i=1}^N \left[m_i m_i' \Big|_{\theta_0} \right]$$

In practice, G_0 and S_0 are estimated by evaluating the analytical expressions at the GMM estimate, $\hat{\theta}_{GMM}$. In the remainder of the paper we refer to the standard errors derived from (4) as the analytical standard errors.

In recent times there has been a growing interest in problems of weak identification in econometric models.² Weak identification occurs when the moment condition is not zero but still very small at parameter values other than θ_0 . This gives rise to objective functions with ridges or near flat spots in the region of the true parameter vector, θ_0 . Stock and Wright (2000) consider the GMM estimator and show that the asymptotic theory devised for identified models does not provide a good approximation to the finite sample distribution of the parameters in a weakly identified model, even for very large but finite sample sizes. They also show that the GMM estimator for a weakly identified parameter is inconsistent and the analytical

² For a review see Stock et al (2002).

standard errors do not converge to the true standard errors. Moreover, they demonstrate that while the strongly identified parameters are consistently estimated, the normal distribution is unlikely to be a good approximation to the finite sample distribution for these parameters. Thus, the presence of some weakly identified parameters can lead to incorrect inferences on well-identified parameters, even in very large samples. They illustrate their findings using both a simultaneous equation model and a consumption based asset pricing model. In our paper, we consider identification in a different context, one based on the use of panel data to estimate the covariance structure of earnings.

3. The GMM Approach to Estimating Earnings Covariance Structures

When considering the covariance structure of earnings, the standard approach is to write earnings as the sum of a permanent component, due to fixed characteristics such as the level of education, and a transitory one, reflecting temporary shocks that affect the individual or the labour market. The objective is to measure the separate roles played by the permanent and transitory shocks in determining inequality and to examine how this may have changed over time. Formally, earnings, y_{it} , are written as

$$y_{it} = p_t \alpha_i + \lambda_t v_{it} \quad (5a)$$

where α_i (the permanent component) and v_{it} (the transitory component) are random variables with means zero and variances σ_α^2 and σ_{vt}^2 respectively; p_t and λ_t are ‘factor loadings’ that allow these variances to change over time in a way that is common across individuals. The inclusion of these factor loadings was an important innovation in models of earnings dynamics and allowed for the existence of structural shifts by

calendar time in the earnings process (Moffitt and Gottschalk (1995)). To capture persistence in the transitory shock, we model v_{it} as an AR(1) process so that

$$v_{it} = \rho v_{it-1} + \varepsilon_{it}. \quad (5b)$$

Other, more elaborate specifications are often used, such as modelling the transitory shock using an ARMA rather than an AR process, allowing for individual heterogeneity and/or a random walk element in the permanent component, including cohort effects in the permanent and/or transitory components and allowing the initial variances (σ_{v1}^2) to vary by age cohort. However, the relatively simple model in (5a) and (5b) captures most of the important features of earnings dynamics, namely time-varying parameters and serial correlation of the transitory shocks. Furthermore, and as discussed in more detail later, the key issues that we raise in this paper turn out to be central to the identification of more elaborate models as well, so that the lessons learned from our analysis carry over to the more complicated models .

The model is estimated by GMM, whereby sample moments are matched to population moments. In this specification, the true variance-covariance matrix has diagonal elements:

$$\sigma_1^2 = p_1^2 \sigma_\alpha^2 + \lambda_1^2 \sigma_{v1}^2, \text{ for } t = 1$$

$$\sigma_t^2 = p_t^2 \sigma_\alpha^2 + \lambda_t^2 (\rho^{2t-2} \sigma_{v1}^2 + \sigma_\varepsilon^2 \sum_{w=0}^{t-2} \rho^{2w}), \text{ for } t > 1$$

and off-diagonal elements:

$$Cov(y_t, y_{t+s}) = p_t p_{t+s} \sigma_\alpha^2 + \lambda_t \lambda_{t+s} (\rho^s \sigma_{v1}^2), \text{ for } t = 1, s > 0$$

$$Cov(y_t, y_{t+s}) = p_t p_{t+s} \sigma_\alpha^2 + \lambda_t \lambda_{t+s} (\rho^{2t+s-2} \sigma_{v1}^2 + \rho^s \sigma_\varepsilon^2 \sum_{w=0}^{t-2} \rho^{2w}), \text{ for } t > 1, s > 0$$

These moments constitute the elements of the vector $m(Y; \theta)$ referred to in Section 2, where the parameter vector to be estimated is given by $\theta = \{\sigma_\alpha^2, \rho, \sigma_\varepsilon^2, \sigma_{v1}^2, p_1 \dots p_T, \lambda_1 \dots \lambda_T\}$. Identification requires a normalization of the factor loadings; in keeping with the literature, we set λ_1 and p_1 equal to one.

4. Results

4a. Analytical Results

In this section we examine the identification issues that arise in GMM estimation of the parameters of the DGP given in (5a) and (5b). The identification problems associated with this model are most easily understood by considering the matrix

$$E \left[\frac{\delta m}{\delta \theta} \Big|_{\theta_0} \right].$$

Using a first order Taylor-series expansion, we can deduce that a

sufficient condition for local identification is that this matrix is of full column rank. If this condition is satisfied then θ_0 is the unique solution to equation (1), within some neighbourhood of θ_0 .

Table 1a provides the typical elements of this matrix for our model, where we initially consider a simplified version without factor loadings and a panel length of 8. The key columns to consider in this case are column (2), the derivative with respect to σ_α^2 and column (4), the derivative with respect to σ_{v1}^2 . The derivative with respect to σ_α^2 is a column of ones while the derivative with respect to σ_{v1}^2 is either one or ρ raised to a power greater than or equal to one. We see from this that the closer ρ is to one, the more similar these two columns are, leading to a failure of the rank condition and problems of identification. In this case only the sum $(\sigma_\alpha^2 + \sigma_{v1}^2)$ may be identified. These derivatives also show how longer panels can assist with identification; more

time periods result in terms containing ρ raised to higher powers, which makes the two derivative vectors increasingly different.

As mentioned earlier, when modelling earnings dynamics it is important to allow for structural changes in the earnings process. Table 1b reproduces the derivative matrix for a model that includes time effects. This allows us to see the crucial role of time trends (and therefore the underlying evolution of inequality) in determining identification of the model. From the discussion in the previous paragraph, we recall that the identification problem associated with high values of ρ arises because of a difficulty in distinguishing between the derivatives with respect to σ_α^2 and σ_{v1}^2 . However, looking at columns (2) and (4) we now see that these derivative vectors include the factor loadings on the permanent and transitory components. If the importance of the transitory component is falling sufficiently quickly relative to the permanent component, this reinforces the rate of decline in the σ_{v1}^2 derivative vector, thus aiding identification. However, the opposite outcome is also possible; a transitory component that is growing more quickly than the permanent component may slow down the rate of decline in the σ_{v1}^2 derivative vector. As a result, a combination of persistence and panel length that allows identification of a model without time trends may not be sufficient to identify a model with time trends.

It is interesting to note that the lessons learned from this analysis carry over to more elaborate models of earnings dynamics. As discussed in Section 3, a number of researchers have estimated models that incorporate heterogeneous growth profiles and consider an ARMA rather than an AR process on the transitory variance shock. However it can be shown that the derivatives with respect to both σ_α^2 and σ_{v1}^2 in a model with heterogeneity and an ARMA error process are identical to those presented

in Table 1b. Therefore it follows that the identification issues that we raise in this paper also apply to these models.

In the next section we provide Monte Carlo evidence of models where identification is facilitated by the presence of time trends in the DGP, as well as instances where identification is hampered by time trends. We also discuss the practical implications of our findings.

4b. Monte Carlo Simulation Results

To examine the implications of the identification problems discussed in the previous section, we conduct a series of Monte Carlo simulations. Multiple data sets are generated using the DGP given by (5a) and (5b), for a range of parameter values, sample sizes and panel lengths.

We first consider the simple model with no time trends. We choose ‘true’ values of $\sigma_\alpha^2 = 0.5$, $\sigma_{v1}^2 = 0.3$ and $\sigma_\varepsilon^2 = 0.2$, in line with results reported in the empirical literature. Initially, we take a very high value of $\rho = 0.95$, a short panel length, $T = 8$ and a sample size of 2,000. This combination of panel length and sample size is chosen to reflect typical sample conditions found in the European Community Household Panel (ECHP).³

Results based on 10,000 simulations are shown in the top panel of Table 2 and Figure 1. The top panel of Figure 1 provides the density of the GMM estimator for the key parameters, while the lower panel reports the associated normal quantile plots for these distributions. Looking first at the estimates of ρ and σ_ε^2 , we see that the GMM

³ The ECHP is a widely-used panel for EU countries for 1994-2001. It is the only panel data set containing earnings data available for some European countries and so is well-suited to cross-country comparisons of earnings dynamics.

estimator is reasonably well behaved, even with short panels; these two parameters are consistently and precisely estimated. In addition the analytical standard errors consistently estimate the true standard deviations. The plots in Figure 1 indicate that the distribution of the estimator for these parameters is well approximated by a normal distribution. Furthermore, the empirical size for a two-tailed test of the hypothesis that the parameter equals its true value is close to its theoretical level. These results indicate that standard asymptotic theory is applicable for these parameters in this case.

However, the problem of distinguishing between permanent and transitory inequality with high values of ρ is evident in Table 2 and Figure 1. In this case, σ_α^2 and σ_{v1}^2 are inconsistently estimated with high standard errors. For each of these two parameters, the analytical standard error underestimates the true standard deviation. Furthermore, Figure 1 shows that the empirical distributions of the parameter estimators are skewed and deviate significantly from normality. The failure of normality and the inconsistency of the analytical standard errors for σ_α^2 and σ_{v1}^2 combine to yield some size distortions in standard t -tests based on equation (3b), particularly for one-sided tests. For example the empirical size of a one-sided theoretical 5% test that $\sigma_\alpha^2 > 0.5$ against $\sigma_\alpha^2 < 0.5$ was almost 0. Finally we note that in almost 2% of cases, the estimated value of σ_α^2 was negative. Reports of negative variances are not unusual in this literature and are often interpreted as a sign that the underlying model is misspecified. However, our analysis shows that negative variances may be a symptom of weak identification rather than model misspecification.

As noted in the analytical section, the identification problem concerning σ_α^2 and σ_{v1}^2 that arises when ρ is high may be overcome by using longer panels. We

consider this issue in the second part of Table 2, which shows the results when we increase T to 25. This is the panel length used by Haider (2001) and is typical of datasets such as the British Household Panel Survey (BHPS) and the Panel Study of Income Dynamics (PSID), which currently have panels of 17 and 36 years respectively. Figure 2 shows the probability density function of the parameters and the associated normal quantile plots for this case. We see that the problems associated with σ_α^2 and σ_{v1}^2 are significantly reduced when we use the longer panel; the parameters are precisely estimated, the analytical standard errors consistently estimate the true values and the parameter estimators are better approximated by a normal distribution.

A question that often arises in this literature concerns the level of ρ at which identification becomes a problem. For the simple model analysed thus far, we can shed some light on this question. The estimates for our model when $\rho = 0.8$ are given in Table 3 and Figure 3. The results show that even with 2,000 observations and a panel of only 8 years, the model is well identified. All the estimated parameters are close to the truth on average and are precisely estimated. In addition the analytical standard errors are unbiased estimators of the truth and the empirical sizes based on the normal approximation are in line with theoretical predictions. Further analysis suggests that serious problems of identification only emerge in this model for values of ρ at or above 0.95

In the remainder of this section we provide Monte Carlo evidence on the identification of models that allow for trends in the permanent and transitory variances. In the analytical section, we showed that the effect of time trends on the identification of the model using GMM is ambiguous. We noted that there may be

models that can be identified using short panels, even though persistence is high ($\rho = 0.95$). For this to happen, the evolution of the transitory component has to be such that it falls sufficiently fast relative to the permanent component. In Table 4 and Figure 4, we present results for a DGP in which p_t increases by 0.01 and λ_t declines by 0.12 in each successive period. In contrast to the model with no time trends, we see that the GMM estimator does a reasonable job of identifying the underlying model, even though persistence is very high. Both the parameter estimates and the analytical standard errors are closer to the truth on average. Furthermore there is a significant improvement in the precision of σ_α^2 and $\sigma_{v_1}^2$, and their empirical distributions are well approximated by normal distributions.

However in Section 3 we also noted that some combinations of persistence and panel length that allow identification of a model without time trends may not be sufficient to identify a model with time trends. Table 5 and Figure 5 reports results for a lower level of persistence ($\rho = 0.8$), but with the transitory component growing faster than the permanent component.⁴ In particular we allow the p_t to increase by 0.01 and the λ_t to rise by 0.03 over the eight year period. It is clear from the results that this model is not well identified by the GMM estimator. Firstly the model failed to converge in almost 1% of the simulations. For the samples in which the model did converge, with the exception of ρ , all of the parameters are poorly estimated and the analytical standard errors are poor estimators of the true deviation.⁵ Furthermore, the normal distribution does not provide a satisfactory approximation to the empirical distribution, with some evidence of bimodality in the distributions of σ_α^2 and $\sigma_{v_1}^2$. In

⁴ This is a feature of many of the new theories of growth, e.g. Violante (2002).

⁵ Because the distributions of the estimated analytical errors are highly skewed for this model, Table 5 reports both the mean and the median of the estimated analytical standard errors.

addition there are very large size distortions for the hypothesis considered. The actual size of the nominal 5% hypothesis test ranged from 14% for ρ to 19% for σ_{v1}^2 . We also found that in 1% of the simulations that converged, the estimator resulted in negative values for the transitory factor loadings, which are difficult to interpret in the context of this model. The problem of negative time trends has been encountered in the empirical literature and our results suggest that this may indicate weak identification.

The results provided in the three lower panels of Table 5 and in Figures 6 to 8 examine the sensitivity of our results to sample size. We consider increasing the sample size to 5,000, 10,000 and 40,000 respectively. Sample sizes in typical survey panels such as the ECHP, the BHPS or the PSID are often significantly less than 5,000, while administrative data sets, such as those used by Dickens (2000), Baker and Solon (2003), Gustavsson (2004) and Capellari (2004), can provide sample sizes of the order of 40,000 or larger. We see that even for sample sizes as large as 10,000, the results are disappointing. Some of the parameters are imprecisely estimated with analytical standard errors that are biased. Furthermore, there is still evidence of significant size distortions when conducting hypothesis tests, as well as significant deviations from normality in the distributions of the estimators. If one has access to sample sizes of the order of 40,000, then both the parameters and standard errors are consistently estimated. However, even with this very large sample size, the empirical distributions shown in Figure 8 still exhibit deviations from normality for many of the underlying distributions, leading to size distortions when conducting hypothesis tests.

Finally in this section, we consider some practical implications of our results. Our analysis has shown that focusing solely on the persistence of transitory shocks when considering identification of these models is inappropriate. Identification also

depends on the evolution of inequality over time. While some patterns of permanent and transitory inequality can assist identification, others can hinder it. We now consider whether the time trends that cause problems for identification are plausible. Further analysis shows that when ρ is less than or equal to 0.6, the model is well identified using relatively short panels ($T = 8$) for any reasonable set of time trends. The transitory variance would have to increase by approximately 300% relative to the permanent variance over an eight year period before problems of identification would arise. As we are not aware of any empirical study that has reported such dramatic differences in the relative growth rates of permanent and transitory inequality, such trends seem implausible. This establishes a cut-off for the persistence of the transitory shock below which our model is identified with as little as eight years of data. For levels of persistence above this cut-off, identification will be sensitive to the evolution of inequality.

5. Conclusion

In this paper we examine the performance of the GMM estimator in the context of the covariance structure of earnings using a relatively simple model of earnings dynamics that nevertheless captures the important features of the process. We examine the sensitivity of parameter identification to key features such as panel length, number of observations, the degree of persistence of earnings shocks and the evolution of inequality over time.

While traditional analysis in this literature has tended to focus on the degree of transitory earnings persistence when considering identification, we show that this emphasis can be misleading. Identification in these models depends on the interaction of the time pattern of inequality in the economy under consideration with the degree

of persistence. In theory, it is always possible to find some pattern of inequality which will cause problems for the GMM estimator no matter how low persistence is. However, our analysis suggests that provided the value of ρ is below 0.6, the AR model considered here is well identified for any reasonable set of time trends, even if only short panels are available.

References

Altonji, J. and L. Segal (1996), 'Small-Sample Bias in GMM Estimation of Covariance Structures,' *Journal of Business & Economic Statistics*, Vol. 14, No. 3, pp. 353-366.

Baker, M. and G. Solon (2003), 'Earnings Dynamics and Inequality among Canadian Men, 1976-1992: Evidence from Longitudinal Income Tax Records', *Journal of Labor Economics*, Vol. 21, No. 2, pp. 289-321.

Bonhomme, S. and J.-M. Robin (2009), 'Assessing the Equalizing Force of Mobility Using Short Panels: France 1990-2000,' *Review of Economic Studies*, Vol. 76, No. 1, pp. 63-92.

Cameron, A.C and P.K. Trivedi (2005), Microeconometrics: Methods and Applications, New York: Cambridge University Press.

Cappellari, L., (2004), 'The Dynamics and Inequality of Italian Men's Earnings: Long-Term Changes or Transitory Fluctuations?' *Journal of Human Resources*, Vol. 39(2), pp. 475-499.

Cervini, M. and X. Ramos (2008), 'Long-Term Earnings Inequality, Earnings Instability and Temporary Employment in Spain 1993-2000,' IZA Working Paper 3538.

Clark, T. (1996), 'Small-Sample Properties of Estimators of Nonlinear Models of Covariance Structure,' *Journal of Business & Economic Statistics*, Vol. 14, No. 3, pp. 367-373.

Daly, M. and R. Valletta (2008), 'Cross-National Trends in Earnings Inequality and Instability', *Economic Letters*, Vol. 99(2), pp. 215-219.

Dearden, L, E. Fitzsimons, A. Goodman and G. Kaplan (2006) 'Estimating the Lifetime Earnings Distribution using Copulas,' IFS Working paper

Dickens, R. (2000). 'The Evolution of Individual Male Earnings in Great Britain: 1975-95,' *Economic Journal*, Vol. 110, No. 460, pp. 27-49.

Doris, A., D. O'Neill, O. Sweetman (2008), 'Does Growth Affect the Nature of Inequality?: Ireland 1994-2001,' IZA Discussion paper 3701.

Haider, S. (2001), 'Earnings Instability and Earnings Inequality of Males in the United States: 1967-1991', *Journal of Labor Economics*, Vol. 19(4), pp. 799-836.

Hall, A. (2005) Generalised Method of Moments, Oxford: Oxford University Press.

Hansen, L.P. (1982) 'Large Sample Properties of Generalised Method of Moments Estimators,' *Econometrica*, Vol. 50, pp. 1029-1054.

Hansen, L., P. Heaton and A. Yaron (1986), 'Finite Sample Properties of Some Alternative GMM Estimators,' *Journal of Business & Economic Statistics*, Vol. 14, pp. 262-280.

Kalwij, A. and R. Alessie (2007), 'Permanent and Transitory Wages of British Men, 1975-2001: Year, Age and Cohort Effects,' *Journal of Applied Econometrics*, Vol. 22, pp. 1063-1093.

Moffitt, R. and P. Gottschalk (1995), 'Trends in the Autocovariance Structure of Earnings in the U.S., 1969-1987,' mimeo, Johns Hopkins University.

Moffitt, R. and P. Gottschalk (2002), 'Trends in the Transitory Variance of Earnings in the United States,' *Economic Journal*, Vol. 112, pp. 68-73.

Moffitt, R. and P. Gottschalk (2008), 'Trends in the Transitory Variance of Male Earnings in the U.S., 1970-2004', Boston College Working Paper.

Newey, W.K, and D. McFadden (1994), 'Large Sample Estimation and Hypothesis Testing,' in Engle and McFadden (eds) *Handbook of Econometrics*, Vol. 4, pp. 2113-2247.

Pozzi, L (2003), 'The Coefficient of Relative Risk Aversion: a Monte Carlo Study Investigating Small Sample Estimator Problems,' *Economic Modelling*, Vol. 20(5), pp. 923-940.

Ramos, X. (2003), 'The Covariance Structure of Earnings in Great Britain, 1991-1999', *Economica*, Vol. 70, pp. 353-374.

Shin, D. and G. Solon (2008), 'Trends in Men's Earnings Volatility: What Does the PSID Show,' mimeo.

Sologon, D. and C. O'Donoghue (2009). 'Earnings Dynamics and Inequality in the EU; 1994-2001,' Working Paper, Maastricht University.

Stock, J., J. Wright and M. Yogo (2002), 'A Survey of Weak Instruments and Weak Identification in Generalised Method of Moments,' *Journal of Business & Economic Statistics*, Vol. 20, No. 4, pp. 518-529.

Stock, J and J. Wright (2000). 'GMM with Weak Identification,' *Econometrica*, Vol. 68, pp. 1055-96.

Tauchén, G. (1986), 'Statistical Properties of GMM Estimators of Structural Parameters Obtained from Financial Market Data,' *Journal of Business Economics and Statistics*, Vol. 4, pp. 397-425.

Violante, G. (2002), 'Technological Acceleration, Skill Transferability and the Rise in Residual Inequality,' *Quarterly Journal of Economics*, Vol. 117(1), pp. 297-338.

Table 1a. Moment Expressions and their Derivatives in Model with No Factor Loadings ($T = 8$)

Moment Expression	$\frac{\partial(\cdot)}{\partial\sigma_\alpha^2}$ (2)	$\frac{\partial(\cdot)}{\partial\rho}$ (3)	$\frac{\partial(\cdot)}{\partial\sigma_{v1}^2}$ (4)	$\frac{\partial(\cdot)}{\partial\sigma_\varepsilon^2}$ (5)
$\sigma_1^2 = \sigma_\alpha^2 + \sigma_{v1}^2$	1	0	1	0
$\sigma_2^2 = \sigma_\alpha^2 + \rho^2\sigma_{v1}^2 + \sigma_\varepsilon^2$	1	$2\rho\sigma_{v1}^2$	ρ^2	1
$\sigma_3^2 = \sigma_\alpha^2 + (\rho^4\sigma_{v1}^2 + \sigma_\varepsilon^2(1 + \rho^2))$	1	$4\rho^3\sigma_{v1}^2$	ρ^4	$1 + \rho^2$
\vdots	\vdots	\vdots	\vdots	\vdots
$\sigma_8^2 = \sigma_\alpha^2 + \rho^{14}\sigma_{v1}^2 + \sigma_\varepsilon^2(1 + \rho^2 + \rho^4 + \dots + \rho^{12})$	1	$14\rho^{13}\sigma_{v1}^2$	ρ^{14}	$1 + \rho^2 + \rho^4 + \dots + \rho^{12}$
$\sigma_{12} = \sigma_\alpha^2 + \rho\sigma_{v1}^2$	1	σ_{v1}^2	ρ	0
$\sigma_{23} = \sigma_\alpha^2 + \rho^3\sigma_{v1}^2 + \rho\sigma_\varepsilon^2$	1	$3\rho^2\sigma_{v1}^2$	ρ^3	ρ
$\sigma_{34} = \sigma_\alpha^2 + \rho^5\sigma_{v1}^2 + \rho\sigma_\varepsilon^2(1 + \rho^2)$	1	$5\rho^4\sigma_{v1}^2$	ρ^5	$\rho + \rho^3$
\vdots	\vdots	\vdots	\vdots	\vdots
$\sigma_{13} = \sigma_\alpha^2 + \rho^2\sigma_{v1}^2$	1	$2\rho\sigma_{v1}^2$	ρ^2	0
$\sigma_{24} = \sigma_\alpha^2 + \rho^4\sigma_{v1}^2 + \rho\sigma_\varepsilon^2$	1	$4\rho^3\sigma_{v1}^2$	ρ^4	ρ^2
$\sigma_{35} = \sigma_\alpha^2 + \rho^6\sigma_{v1}^2 + \rho\sigma_\varepsilon^2(1 + \rho^2)$	1	$6\rho^5\sigma_{v1}^2$	ρ^6	$\rho^2 + \rho^4$
\vdots	\vdots	\vdots	\vdots	\vdots
$\sigma_{18} = \sigma_\alpha^2 + \rho^7\sigma_{v1}^2$	1	$7\rho^6\sigma_{v1}^2$	ρ^7	0

Table 1b. Moment Expressions and their Derivatives in Model with Factor Loadings ($T = 8$)

Moment Expression	$\frac{\partial(\cdot)}{\partial\sigma_\alpha^2}$	$\frac{\partial(\cdot)}{\partial\rho}$	$\frac{\partial(\cdot)}{\partial\sigma_{v1}^2}$	$\frac{\partial(\cdot)}{\partial\sigma_\varepsilon^2}$
$\sigma_1^2 = p_1^2\sigma_\alpha^2 + \lambda_1^2\sigma_{v1}^2$	p_1^2	0	λ_1^2	0
$\sigma_2^2 = p_2^2\sigma_\alpha^2 + \lambda_2^2(\rho^2\sigma_{v1}^2 + \sigma_\varepsilon^2)$	p_2^2	$2\rho\lambda_2^2\sigma_{v1}^2$	$\lambda_2^2\rho^2$	λ_2^2
$\sigma_3^2 = p_2^2\sigma_\alpha^2 + \lambda_3^2(\rho^4\sigma_{v1}^2 + \sigma_\varepsilon^2(1 + \rho^2))$	p_3^2	$4\rho^3\lambda_3^2\sigma_{v1}^2$	$\lambda_3^2\rho^4$	$\lambda_3^2(1 + \rho^2)$
\vdots	\vdots	\vdots	\vdots	\vdots
$\sigma_8^2 = p_8^2\sigma_\alpha^2 + \lambda_8^2(\rho^{14}\sigma_{v1}^2 + \sigma_\varepsilon^2(1 + \rho^2 + \rho^4 + \dots + \rho^{12}))$	p_8^2	$14\rho^{13}\lambda_8^2\sigma_{v1}^2$	$\lambda_8^2\rho^{14}$	$\lambda_8^2(1 + \rho^2 + \rho^4 + \dots + \rho^{12})$
$\sigma_{12} = p_1p_2\sigma_\alpha^2 + \lambda_1\lambda_2(\rho\sigma_{v1}^2)$	p_1p_2	$\lambda_1\lambda_2\sigma_{v1}^2$	$\lambda_1\lambda_2\rho$	0
$\sigma_{23} = p_2p_3\sigma_\alpha^2 + \lambda_2\lambda_3(\rho^3\sigma_{v1}^2 + \rho\sigma_\varepsilon^2)$	p_2p_3	$3\lambda_2\lambda_3\rho^2\sigma_{v1}^2$	$\lambda_2\lambda_3\rho^3$	$\lambda_2\lambda_3\rho$
$\sigma_{34} = p_3p_4\sigma_\alpha^2 + \lambda_3\lambda_4(\rho^5\sigma_{v1}^2 + \rho\sigma_\varepsilon^2(1 + \rho^2))$	p_3p_4	$5\lambda_3\lambda_4\rho^4\sigma_{v1}^2$	$\lambda_3\lambda_4\rho^5$	$\lambda_3\lambda_4(\rho + \rho^3)$
\vdots	\vdots	\vdots	\vdots	\vdots
$\sigma_{13} = p_1p_3\sigma_\alpha^2 + \lambda_1\lambda_3(\rho^2\sigma_{v1}^2)$	p_1p_3	$2\lambda_1\lambda_3\rho\sigma_{v1}^2$	$\lambda_1\lambda_3\rho^2$	0
$\sigma_{24} = p_2p_4\sigma_\alpha^2 + \lambda_2\lambda_4(\rho^4\sigma_{v1}^2 + \rho\sigma_\varepsilon^2)$	p_2p_4	$4\lambda_2\lambda_4\rho^3\sigma_{v1}^2$	$\lambda_2\lambda_4\rho^4$	$\lambda_2\lambda_4\rho^2$
$\sigma_{35} = p_3p_5\sigma_\alpha^2 + \lambda_3\lambda_5(\rho^6\sigma_{v1}^2 + \rho\sigma_\varepsilon^2(1 + \rho^2))$	p_3p_5	$6\lambda_3\lambda_5\rho^5\sigma_{v1}^2$	$\lambda_3\lambda_5\rho^6$	$\lambda_3\lambda_5(\rho^2 + \rho^4)$
\vdots	\vdots	\vdots	\vdots	\vdots
$\sigma_{18} = p_1p_8\sigma_\alpha^2 + \lambda_1\lambda_8(\rho^7\sigma_{v1}^2)$	p_1p_8	$7\lambda_1\lambda_8\rho^6\sigma_{v1}^2$	$\lambda_1\lambda_8\rho^7$	0

Table 2. Monte Carlo Simulations for DGP with Very High Persistence of Transitory Shock and No Factor Loadings. 10,000 Replications.

True Parameter Values: $\rho = 0.95$, $\sigma_\alpha^2 = 0.5$, $\sigma_\varepsilon^2 = 0.2$, $\sigma_{v1}^2 = 0.3$. p_t and λ_t equal to one in each period.

Parameter	Mean	Median	Standard Deviation	Mean Analytical Standard Error	Empirical Size (Nominal .025 in both tails)
$T = 8, N = 2,000$					
ρ	0.9501	0.9499	0.0133	0.0133	0.0217 0.0278
σ_α^2	0.4579	0.4978	0.3840	0.2539	0.0000 0.0508
σ_ε^2	0.1999	0.2000	0.0100	0.0099	0.0294 0.0194
σ_{v1}^2	0.3415	0.3011	0.3834	0.2524	0.0563 0.0000
$T = 25, N = 2,000$					
ρ	0.9500	0.9500	0.0045	0.0045	0.0202 0.0309
σ_α^2	0.4981	0.5006	0.0634	0.0636	0.0185 0.0275
σ_ε^2	0.1999	0.1998	0.0088	0.0087	0.0329 0.0191
σ_{v1}^2	0.3014	0.2990	0.0683	0.0687	0.0320 0.0166

Table 3: Monte Carlo Simulations for DGP with Moderately High Persistence of Transitory Shock and No Factor Loadings. 10,000 Replications.

True Parameters: $\rho = 0.8$, $\sigma_\alpha^2 = 0.5$, $\sigma_\varepsilon^2 = 0.2$, $\sigma_{v1}^2 = 0.3$. p_t and λ_t equal to one in each period.

Parameter	Mean	Median	Standard Deviation	Mean Analytical Standard Error	Empirical Size (Nominal .025 in both tails)
$T = 8, N = 2,000$					
ρ	0.800	0.801	0.016	0.0159	0.017 0.037
σ_α^2	0.498	0.50	0.0299	0.0297	0.026 0.023
σ_ε^2	0.1996	0.1999	0.0058	0.0057	0.032 0.019
σ_{v1}^2	0.302	0.301	0.0246	0.0243	0.028 0.025

Table 4: Monte Carlo Simulations for DGP with Very High Persistence of Transitory Error and Factor Loadings. 10,000 Replications.

True Parameters: $\rho = 0.95$, $\sigma_\alpha^2 = 0.5$, $\sigma_\varepsilon^2 = 0.2$, $\sigma_{v1}^2 = 0.3$. p_t increasing by 0.01 and λ_t decreasing by 0.12 in successive periods.

Parameter	Mean	Median	Standard Deviation	Mean Analytical Standard Error	Empirical Size (Nominal .025 in both tails)
$T = 8, N = 2,000$					
ρ	0.9505	0.9510	0.0284	0.0270	0.0358 0.0479
σ_α^2	0.4995	0.4995	0.0272	0.273	0.0263 0.0223
σ_ε^2	0.1991	0.1987	0.0132	0.0133	0.0399 0.0147
σ_{v1}^2	0.3008	0.2996	0.0167	0.0164	0.0443 0.0116

Table 5: Monte Carlo Simulations for DGP with Moderately High Persistence of Transitory Error and Factor Loadings. 10,000 Replications.

True Parameters: $\rho = 0.8$, $\sigma_\alpha^2 = 0.5$, $\sigma_\varepsilon^2 = 0.2$, $\sigma_{v1}^2 = 0.3$. p_t increasing by 0.01 and λ_t increasing by 0.03 in successive periods.

Parameter	Mean	Median	Standard Deviation	Mean Analytical Standard Error (Median)	Empirical Size (Nominal .025 in both tails)
<i>T</i> = 8, <i>N</i> = 2,000					
ρ	0.7940	0.7949	0.0242	0.0232 (0.0243)	0.0453 0.0890
σ_α^2	0.5652	0.4874	0.5405	0.6929 (0.0765)	0.1109 0.0614
σ_ε^2	1.3715	0.1926	16.2046	255.70 (0.0282)	0.1100 0.0096
σ_{v1}^2	0.2345	0.3140	0.5395	0.6898 (0.0736)	0.0515 0.1357
<i>T</i> = 8, <i>N</i> = 5,000					
ρ	0.7970	0.7974	0.0167	0.0162 (0.0164)	0.0418 0.0490
σ_α^2	0.5170	0.4938	0.1543	0.0970 (0.0577)	0.092 0.0245
σ_ε^2	0.387	0.1973	3.3048	24.306 (0.0192)	0.085 0.002
σ_{v1}^2	0.283	0.3067	0.3427	0.0954 (0.056)	0.022 0.102
<i>T</i> = 8, <i>N</i> = 10,000					
ρ	0.7989	0.7989	0.0121	0.0120 (0.0120)	0.036 0.0249
σ_α^2	0.5088	0.4970	0.0823	0.0556 (0.0443)	0.074 0.006
σ_ε^2	0.2268	0.1984	0.7987	0.2329 (0.0139)	0.067 0.004
σ_{v1}^2	0.2913	0.3034	0.0814	0.0544 (0.0432)	0.006 0.078
<i>T</i> = 8, <i>N</i> = 40,000					
ρ	0.7996	0.7996	0.0062	0.0062 (0.0062)	0.033 0.019
σ_α^2	0.5016	0.4990	0.0248	0.0242 (0.0231)	0.0572 0.0016
σ_ε^2	0.2002	0.1997	0.0077	0.0076 (0.0072)	0.0479 0.0035
σ_{v1}^2	0.2984	0.3012	0.0243	0.0238 (0.0026)	0.0014 0.0590

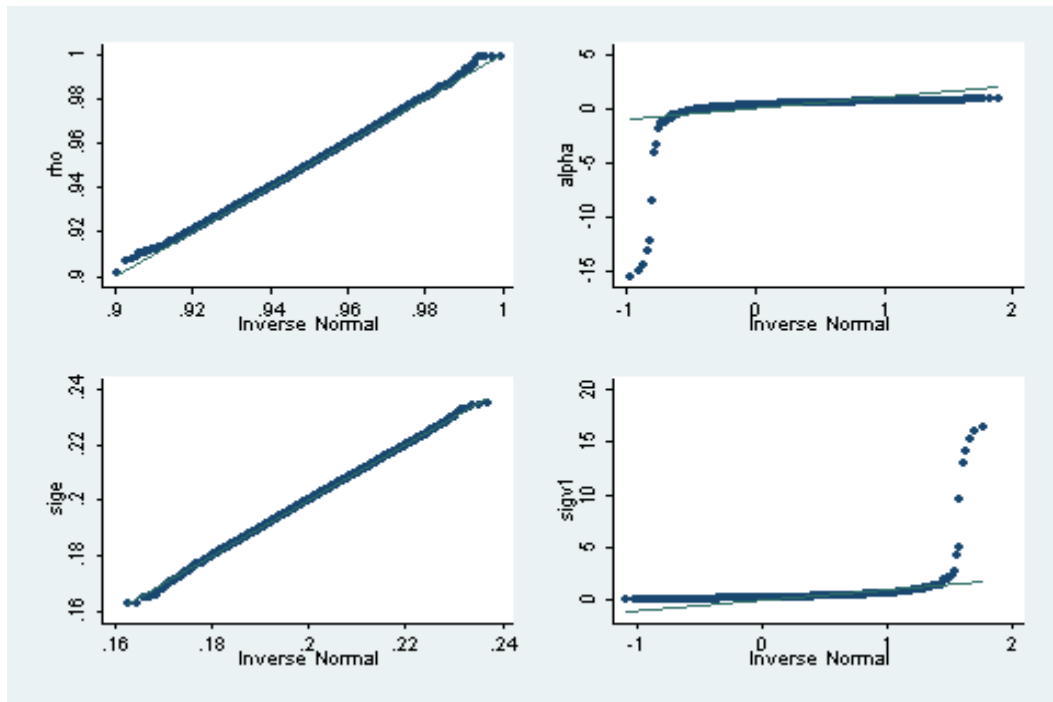
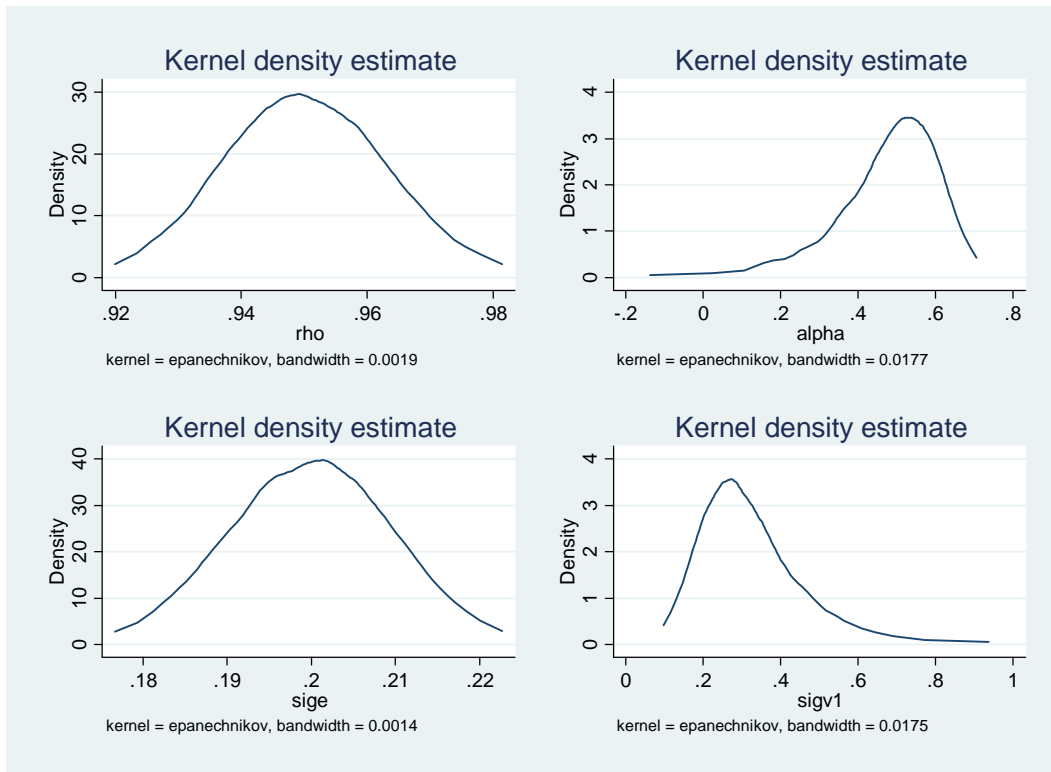


Figure 1: pdf of GMM Estimators (Top Panel) and Normal Quantile Plots (Bottom Panel).

True Parameters: $\rho = 0.95$, $\sigma_\alpha^2 = 0.5$, $\sigma_\varepsilon^2 = 0.2$, $\sigma_{v1}^2 = 0.3$. p_t and λ_t equal to one in each period. $T = 8$, $N = 2,000$.

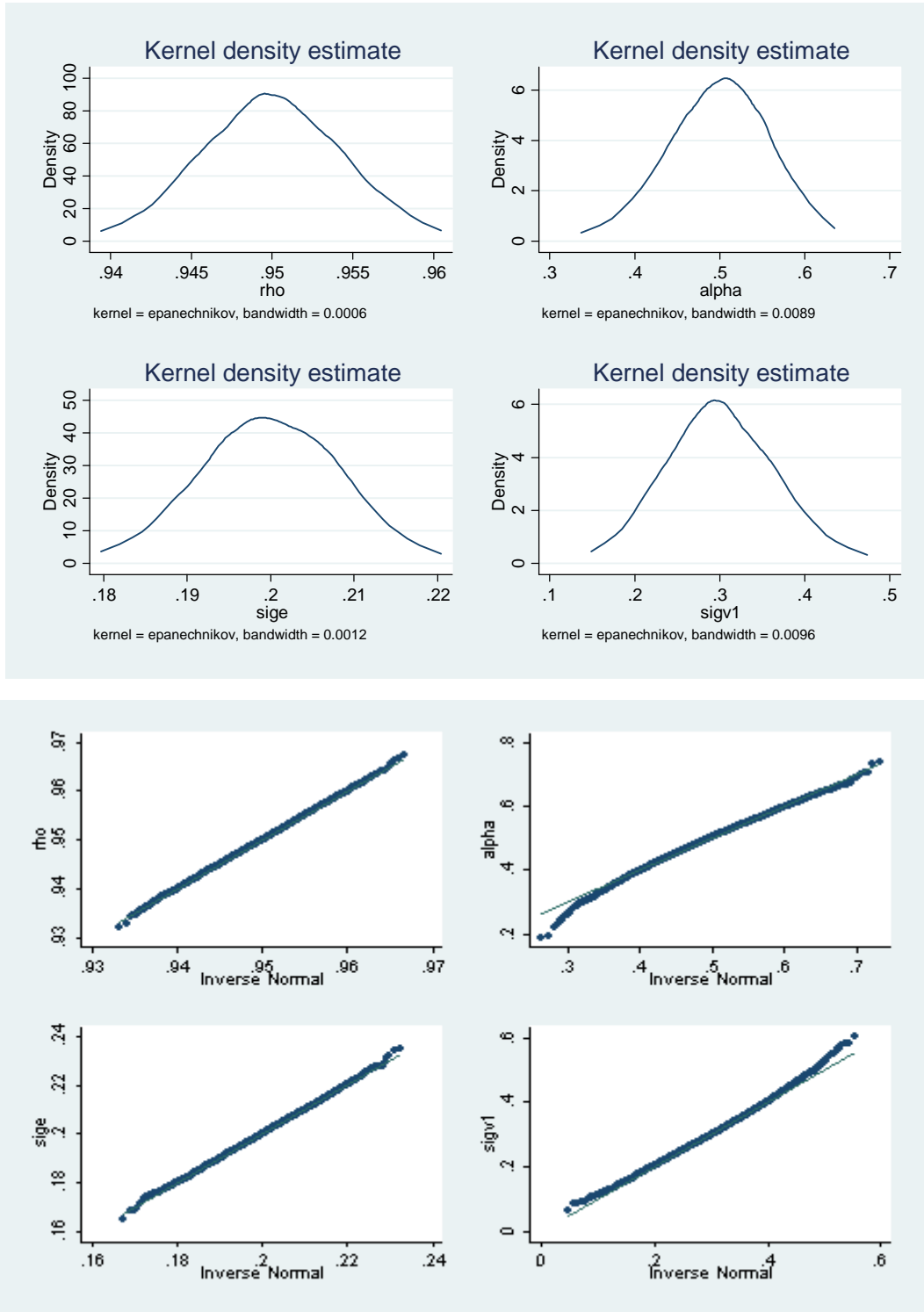


Figure 2: pdf of GMM Estimators (Top Panel) and Normal Quantile Plots (Bottom Panel).

True Parameters: $\rho = 0.95$, $\sigma_\alpha^2 = 0.5$, $\sigma_\varepsilon^2 = 0.2$, $\sigma_{v1}^2 = 0.3$. p_t and λ_t equal to one in each period. $T = 25$, $N = 2,000$.

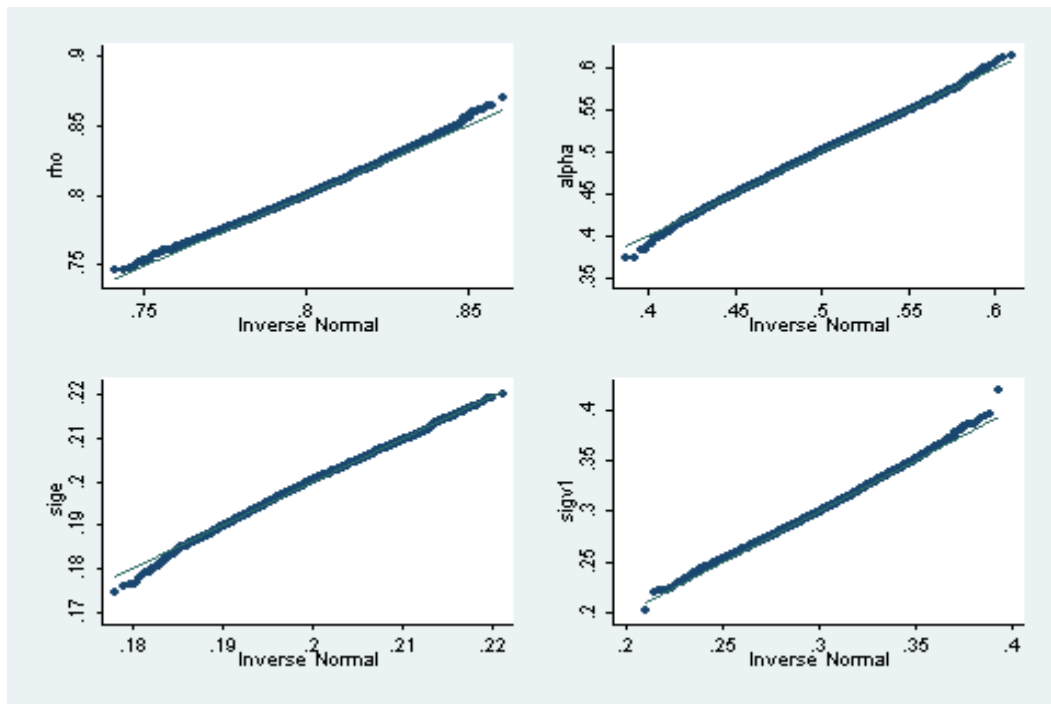
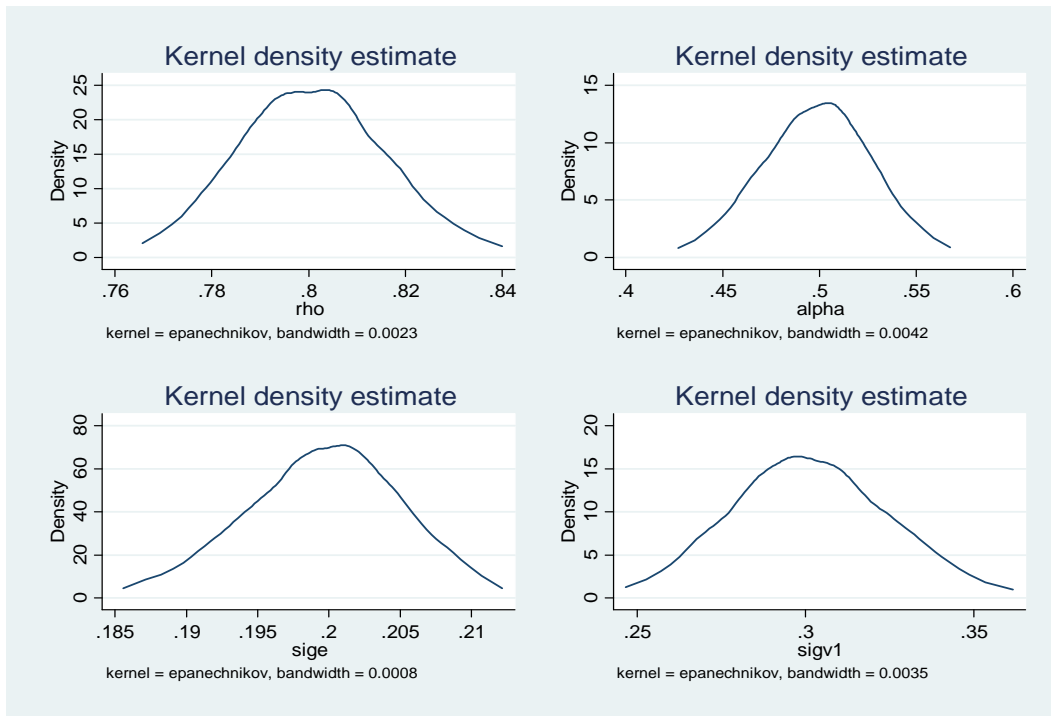


Figure 3: pdf of GMM Estimators (Top Panel) and Normal Quantile Plots (Bottom Panel).

True Parameters: $\rho = 0.8$, $\sigma_\alpha^2 = 0.5$, $\sigma_\varepsilon^2 = 0.2$, $\sigma_{v1}^2 = 0.3$. p_t and λ_t equal to one in each period. $T = 8$, $N = 2,000$.

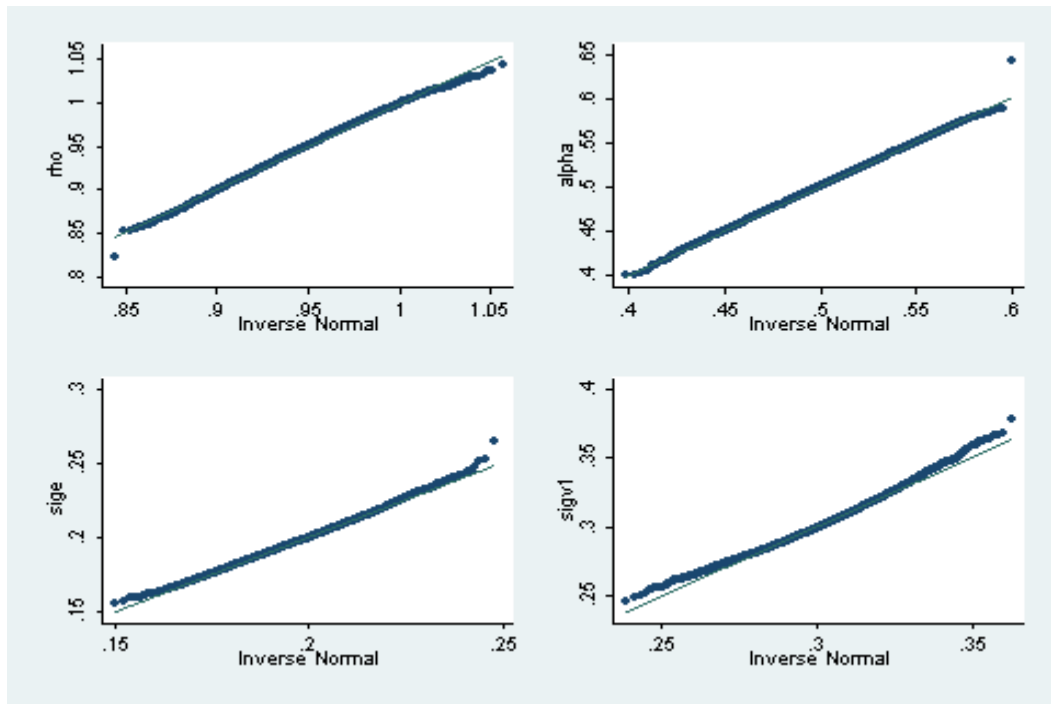
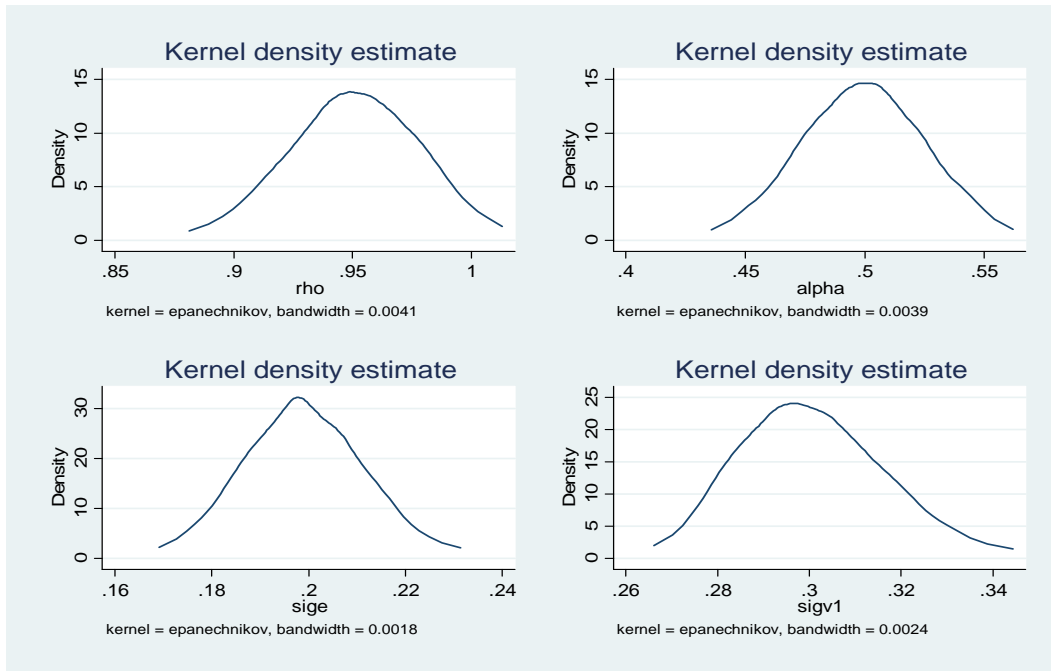


Figure 4: pdf of GMM Estimators (Top Panel) and Normal Quantile Plots (Bottom Panel).

True Parameters: $\rho = 0.95$, $\sigma_\alpha^2 = 0.5$, $\sigma_\varepsilon^2 = 0.2$, $\sigma_{v1}^2 = 0.3$. p_t increasing by 0.01 and λ_t decreasing by 0.12 in successive periods. $T = 8$, $N = 2,000$.

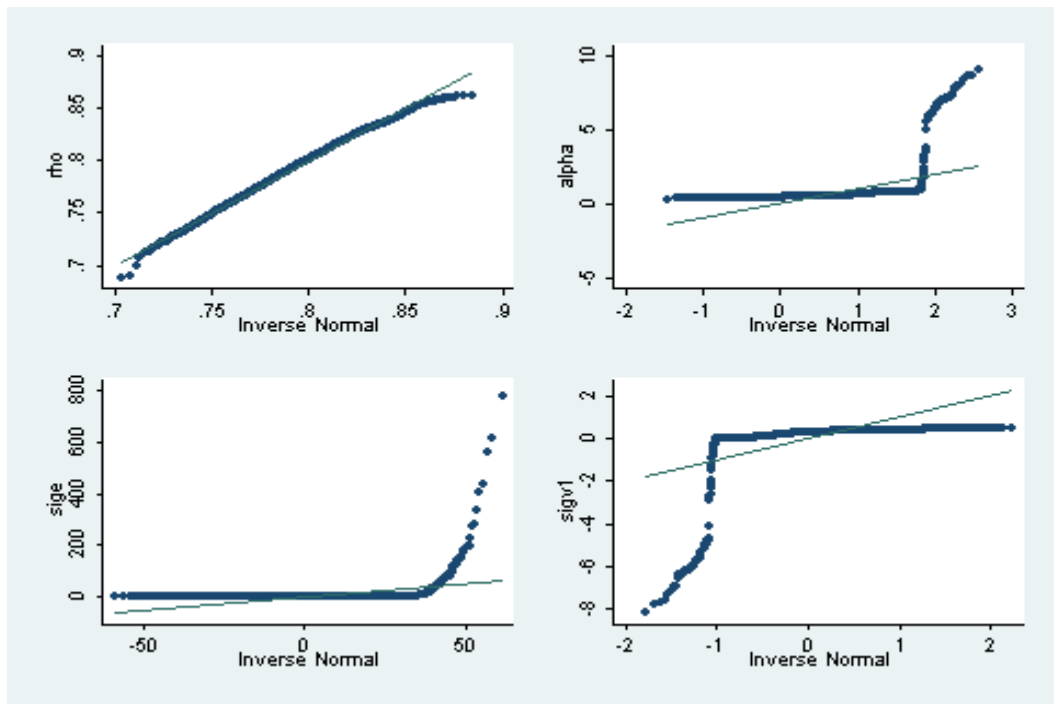
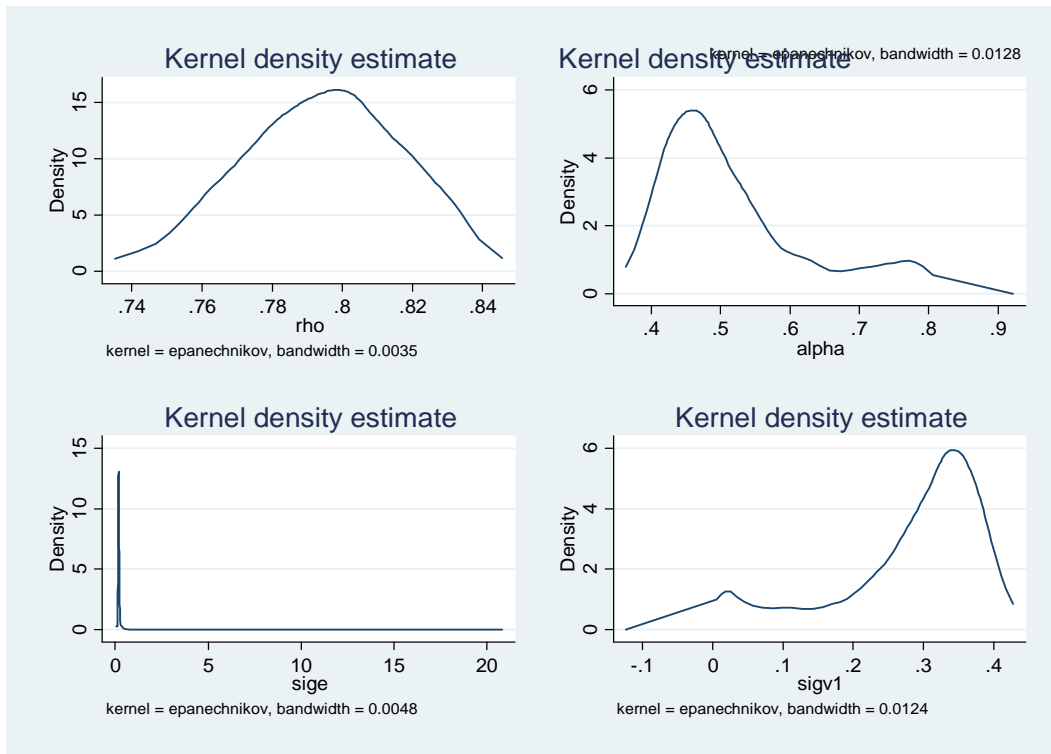


Figure 5: pdf of GMM Estimators (Top Panel) and Normal Quantile Plots (Bottom Panel).

True Parameters: $\rho = 0.8$, $\sigma_\alpha^2 = 0.5$, $\sigma_\varepsilon^2 = 0.2$, $\sigma_{v1}^2 = 0.3$. p_t increasing by 0.01 and λ_t increasing by 0.03 in successive periods. $T = 8$, $N = 2,000$.

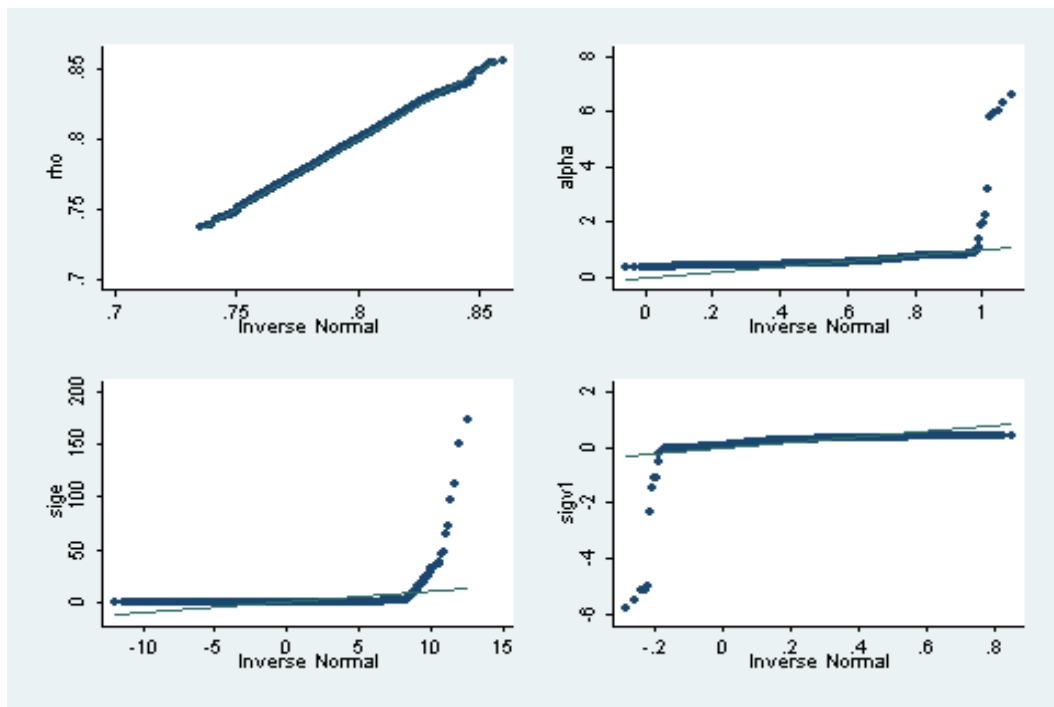
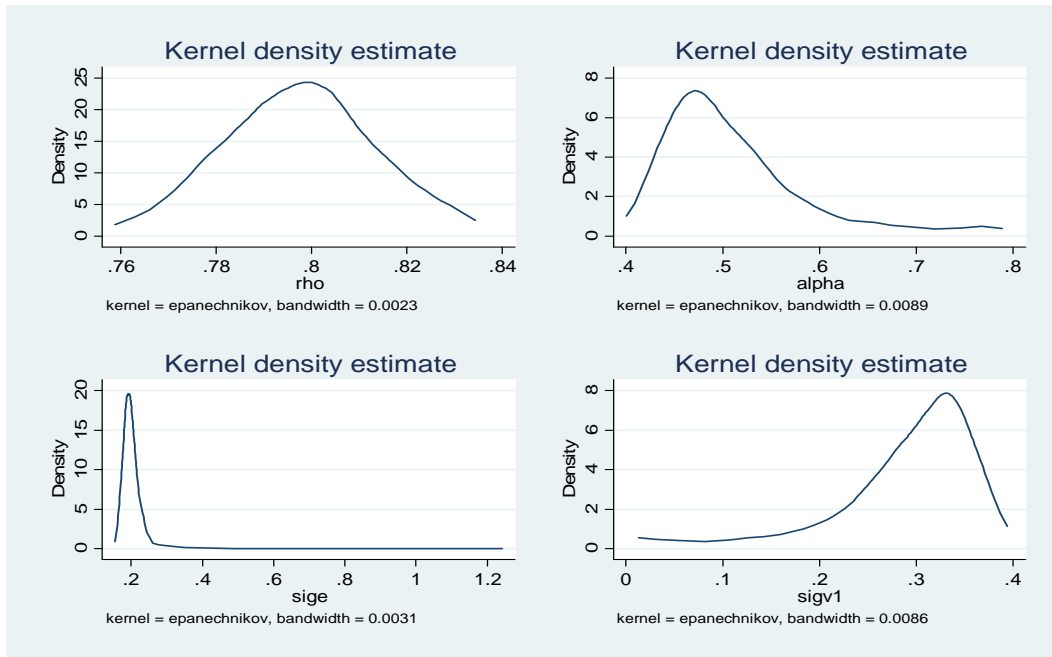


Figure 6: pdf of GMM Estimators (Top Panel) and Normal Quantile Plots (Bottom Panel).

True Parameters: $\rho = 0.8$, $\sigma_\alpha^2 = 0.5$, $\sigma_\varepsilon^2 = 0.2$, $\sigma_{v_1}^2 = 0.3$. p_t increasing by 0.01 and λ_t increasing by 0.03 in successive periods. $T = 8$, $N = 5,000$.

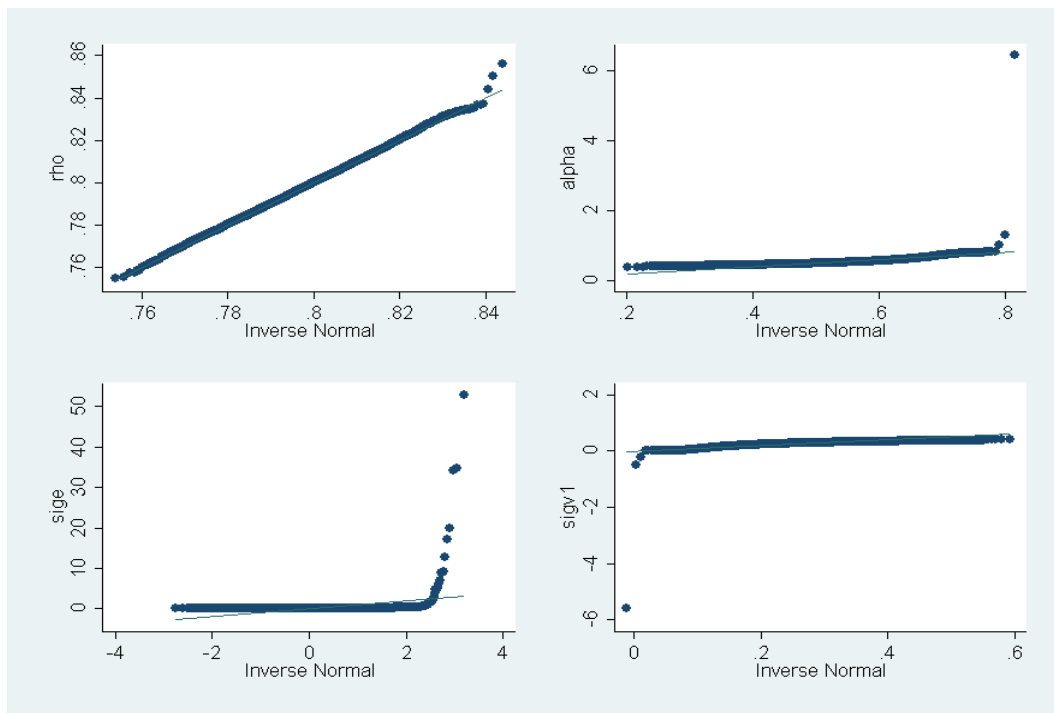
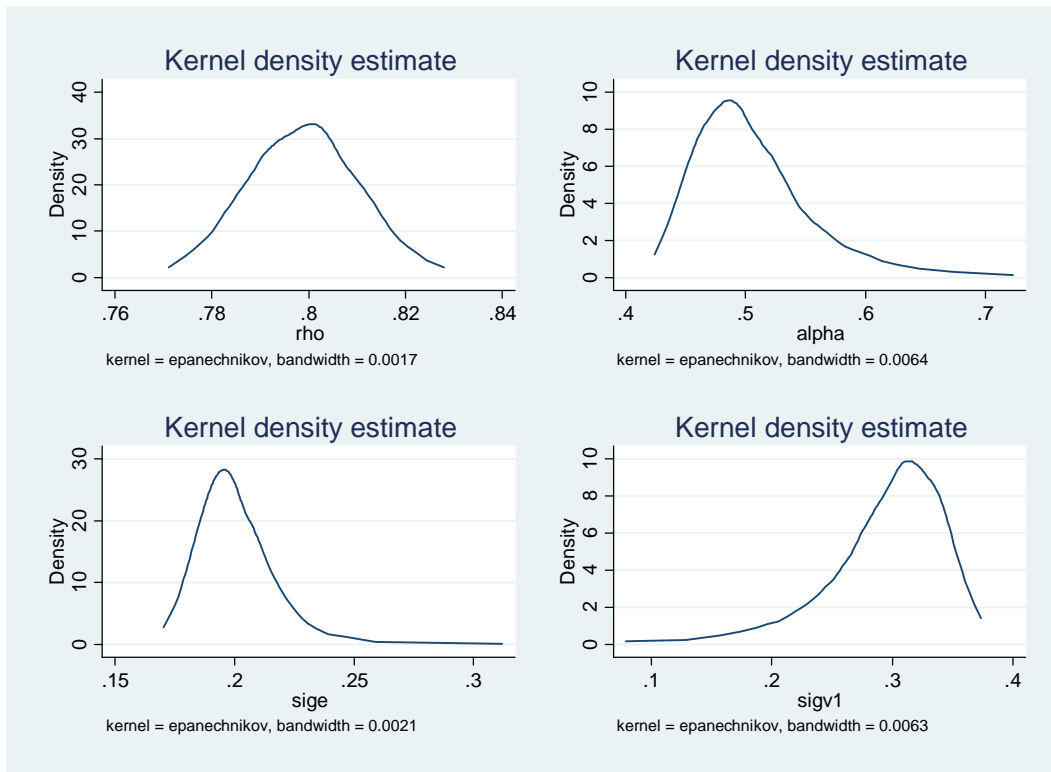


Figure 7: pdf of GMM Estimators (Top Panel) and Normal Quantile Plots (Bottom Panel).

True Parameters: $\rho = 0.8$, $\sigma_\alpha^2 = 0.5$, $\sigma_\varepsilon^2 = 0.2$, $\sigma_{v1}^2 = 0.3$. p_t increasing by 0.01 and λ_t increasing by 0.03 in successive periods. $T = 8$, $N = 10,000$.

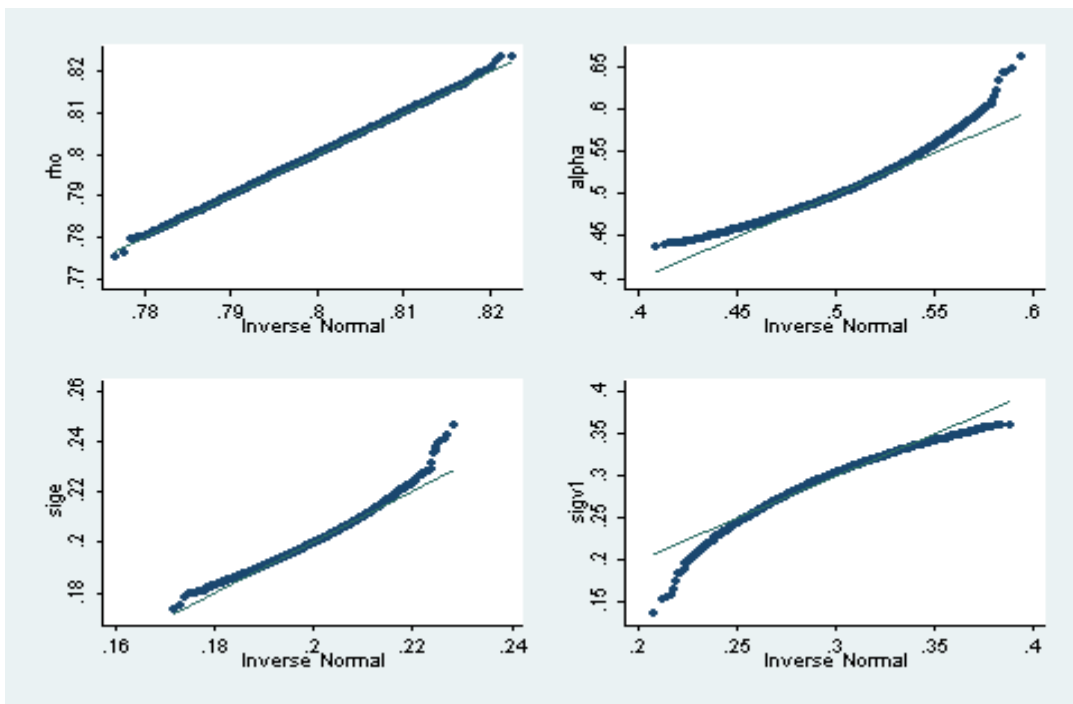
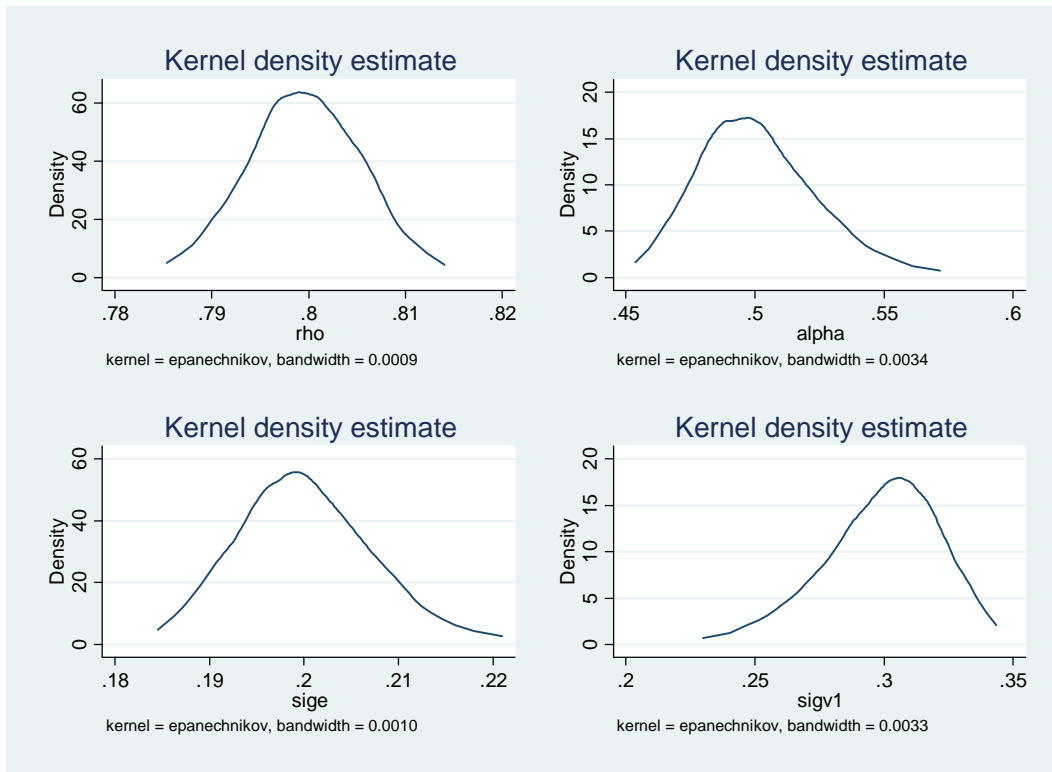


Figure 8: pdf of GMM Estimators (Top Panel) and Normal Quantile Plots (Bottom Panel).

True Parameters: $\rho = 0.8$, $\sigma_\alpha^2 = 0.5$, $\sigma_\varepsilon^2 = 0.2$, $\sigma_{v1}^2 = 0.3$. p_t increasing by 0.01 and λ_t increasing by 0.03 in successive periods. $T = 8$, $N = 40,000$.