Testing for Uncorrelated Residuals in Dynamic Count Models with an Application to Corporate Bankruptcy

Sant'Anna, Pedro H. C.

Universidad Carlos III de Madrid

May 2013

Testing for Uncorrelated Residuals in Dynamic Count Models with an Application to Corporate Bankruptcy

Pedro H. C. Sant'Anna

Universidad Carlos III de Madrid, Getafe, 28903, Spain

pde@eco.uc3m.es

May 2013

Abstract

This article proposes a new diagnostic test for dynamic count models, which is well suited for risk management. Our test proposal is of the Portmanteau-type test for lack of residual autocorrelation. Unlike previous proposals, the resulting test statistic is asymptotically pivotal when innovations are uncorrelated, but not necessarily *iid* nor a martingale difference. Moreover, the proposed test is able to detect local alternatives converging to the null at the parametric rate $T^{-1/2}$, with T the sample size. The finite sample performance of the test statistic is examined by means of a Monte Carlo experiment. Finally, using a dataset on U.S. corporate bankruptcies, we apply our test proposal to check if common risk models are correctly specified.

Keywords: Time Series of counts; Residual autocorrelation function; Model checking; Credit risk management.

1. Introduction

Credit risk affects virtually every financial contract. Therefore the measurement, pricing and management of credit risk have received much attention from economists, bank supervisors and regulators, and financial market practitioners.

A widely used measure of credit risk is the probability of corporate default (PD). Many default risk models that are employed day-to-day on risk management, such as CreditMetrics, Moody's KMV, and CreditRisk+, rely on the assumption of conditionally independent defaults, that is, conditional on observable macroeconomic and financial variables, together with firm specific characteristics, defaults are time independent. Nonetheless, recent studies have found evidence of violation of this assumption, see e.g. Das et al. (2007), Koopman et al. (2011, 2012).

In order to accommodate deviations from conditional independence, richer classes of models have been proposed. Koopman et al. (2011, 2012) consider that a common frailty effect, modeled as a Gaussian AR(1), drives the excess default counts clustering. However, an important question remain unanswered: Is the AR(1) latent process structure enough to capture all the excess default? If this is not the case, there would be evidence of residual serial correlation. Answering this question is appealing for risk management because as shown by of Duffie et al. (2009) and Koopman et al. (2011, 2012), model misspecification may lead to a downward bias when assessing the probability of extreme default losses.

In this paper, we consider a general model check which is well suited to evaluate the correct specification of aggregate default and bankruptcy count models. We propose a new test for serial correlation of multiplicative residuals in a dynamic count data model under weak assumptions, namely when no parametric distribution restrictions are made and the innovations are neither restricted to be *iid* nor a martingale difference. Our test statistic is of the Portmanteau class, and takes a quadratic form in linear combinations of residuals sample autocorrelations. A major advantage of our test statistic is that it is asymptotically distribution-free in the presence of estimated parameters, even when the innovations are not *iid*, which is in contrast with classical lack of autocorrelation tests, e.g. Box and Pierce

(1970) and Ljung and Box (1978). Moreover, the proposed test is able to detect local alternatives converging to the null at the parametric rate $T^{-1/2}$, with T the sample size.

Although the study of conditions for stationarity and ergodicity, and the related asymptotic properties of parameter estimates of a count data models have been an active area of research, see e.g. Tjøstheim (2012) and Fokianos (2012) and references therein, less attention has been placed into model checks. Neumann (2011) and Fokianos and Neumann (2013) propose goodness-of-fit test for the intensity parameter of an observation-driven Poisson time series regression. However, the conditions imposed to justify the corresponding inference are rather restrictive and rule out exponential intensity functions, which is the canonical functional form in count models. Jung and Tremayne (2003) and Sun and McCabe (2013) consider score-type tests for lack of serial dependence in the integer autoregressive (INAR) class of models, that is, if there is need to estimate dynamic count models. Nonetheless, the test is not suitable to test if specifications like the INAR(1) captures all the serial dependence. Moreover, the type of INAR process considered by these authors does not allow to include covariates, which limit its applicability in economics contexts. Davis et al. (2000) consider an overdispersed autocorrelated Poisson model, where the overdispersion and autocorrelation in the count variable are driven by a multiplicative log-normal latent process. Their proposed test statistic is a variant of the Box and Pierce (1970) test (hereafter BP) for lack of autocorrelation. Nonetheless, Davis et al. (2000) asymptotic results are derived under the assumption that the latent process is independent of the covariates. In fact, Davis et al. (2000) consider only strict exogenous (deterministic) covariates, a case with not much applied interest. With all these maintained hypothesis, the distribution of the test statistic under the null is derived under the serial independence assumption, a much stronger condition than lack of correlation.

When the innovations are uncorrelated, but not independent, the use of residuals sample autocorrelations,without proper scaling, might not be appropriate to test for lack of autocorrelation of the innovations. The scaling might depend on higher-order serial dependence of the innovations, the model and the estimator used - see Francq et al. (2005). In this article we follow Delgado and Velasco (2011) approach which supply an asymptotically pivotal transform of the residuals sample autocorrelation, which serves a basis for model checking. Given that the residual transform is asymptotically distribution-free, and hence does not rely on estimation methods nor on high-order dependence assumptions, this procedure is well suited for dynamic count data models.

Our approach does not impose that the innovations are serially independent nor independent from the covariates. In fact, one of the contributions of the paper is to show that if innovations are not independent of the covariates, the distribution of residuals sample autocorrelations is not necessarily pivotal. In other words, the BP test or its variants are not asymptotically distribution-free. In order to illustrate the issues of ignoring the estimation effect and/or possible higher order serial dependence, a simulation exercise compares the finite sample properties of our test with the classical Box and Pierce (1970).

Eventually, we apply our goodness-of-fit test procedure to the risk management context. Considering a set of observed macroeconomic and financial variables as covariates, we evaluate the specification of different models for US bankruptcy counts for big public firms, using monthly data from 1985 to 2012. First, we apply our procedure to test the null hypothesis of lack of residual autocorrelation when only macroeconomic and financial variables are used as covariates. Using our proposed test statistic, we reject the null, which may indicate evidence of a frailty effect in the default count data, confirming the finds of Duffie et al. (2009) and Koopman et al. (2011, 2012).

Once one finds evidence of a frailty effect, it is common to introduce a first order autocorrelated latent process into the model - see for instance Koopman et al. (2011, 2012). Following this proposal, we consider the Davis et al. (2003) observation-driven Poisson GARMA model, with an AR(1) or MA(1) term. In order to access if the inclusion of the additional parameters would suffice to capture all the excess default clustering, we apply again our test statistic on the residuals of these augmented models. In both GARMA models, we fail to reject the null of lack of residual autocorrelation, providing some evidence that considering only first order autocorrelation might suffices to capture the linear dynamics of monthly US bankruptcy counts. To the best of our knowledge, we are the first to formally test if count models with AR(1) of MA(1) are able to capture the linear dynamics in a risk management framework.

The rest of the paper is organized as follows: the framework of our test is presented in the next section. In the third section, we introduce the autocorrelation transformation and discuss its asymptotic properties. In Section 4, we apply the transformation to lack of residual autocorrelation testing. In section 5, we discuss the finite sample properties of the proposed test via Monte Carlo Simulations. Then, we illustrate our test with an empirical application for big public corporate bankruptcies and the last section concludes.

2. Testing lack of autocorrelation on dynamic count data models

To introduce the family of count models considered here, let $\{Y_t\}_{t\in\mathbb{Z}}$ be a stationary time series of counts defined on $\mathbb{N}_0 = \{0, 1, 2, \dots\}$, and suppose that for each t, $\mathbf{X}_t$ is a $k \times 1$ vector of predetermined observed covariates, which first component is assumed to be one. A multiplicative error model is assumed to take the form

$$Y_t = \exp\left(\mathbf{X}_t'\boldsymbol{\beta_0}\right)\varepsilon_t, \tag{1}$$

where $\boldsymbol{\beta_0}$ is a $k\times 1$ vector of unknown parameters, $\{\varepsilon_t\}_{t\in\mathbb{Z}}$ is a stationary unobserved process, such that $E(\varepsilon_t) = 1$ and $Cov(\varepsilon_t, \varepsilon_{t-\tau}) = \gamma_{\beta_0}(\tau)$, $\tau \in \mathbb{Z}$, $\gamma_{\beta_0}(\tau)$ denoting the autocovariance of order τ of ε_t. We denote $\lambda_t = \exp\left(\mathbf{X}_t'\boldsymbol{\beta}\right)$ as the (conditional) mean function of the count process.

The focus of our attention is the autocorrelation function of the multiplicative error, $\{\varepsilon_t\}_{t\in\mathbb{Z}}$,

$$Corr(\varepsilon_t, \varepsilon_{t-\tau}) = \rho_{\beta_0}(\tau) = \frac{\gamma_{\beta_0}(\tau)}{\gamma_{\beta_0}(0)}, \tau \in \mathbb{Z}.$$

Given any model that can be written as (1), the purpose is to test the null hypothesis

$$H_0 : \rho_{\beta_0}(\tau) = 0,\ \tau = 1, \dots, m \text{ and some } \beta_0 \in \Theta$$

against the fixed alternative hypothesis

$$H_1^s : \rho_{\beta_0}(\tau) \neq 0 \text{ for any } \tau = 1, \dots, m \text{ and some } \beta_0 \in \Theta$$

for some $m \geq 1$.

Given observations $\{Y_t, \boldsymbol{X}_t\}_{t=1}^{T}$, $\rho_\beta(\tau)$ is estimated by the sample autocorrelation function

$$\hat{\rho}_{T\beta}(\tau) = \frac{\hat{\gamma}_{T\beta}(\tau)}{\hat{\gamma}_{T\beta}(0)}, \tau \in \mathbb{Z} \tag{2}$$

where

$$\hat{\gamma}_{T\beta}(\tau) = \frac{1}{T}\sum_{t=1+\tau}^{T} \frac{(Y_t - \lambda_t)(Y_{t-\tau} - \lambda_{t-\tau})}{\lambda_t \lambda_{t-\tau}}. \tag{3}$$

It is worth mentioning that the vector of covariates does not include lags of Y_t, since, as shown by Zeger and Qaqish (1988), this leads to non-stationary count process unless positive dependence between the count variables is rule out. Nonetheless, one can include particular functions of lags of Y_t as long as the process is stationary. Examples are the log-linear Poisson Autoregression of Fokianos and Tjøstheim (2011), where $\ln(1 + Y_{t-1})$ is included as covariate, and the Generalized ARMA model of Davis et al. (2003), where $(Y_{t-l} - \lambda_{t-l})/\lambda_{t-l},\ l \geq 1$ is included in the covariate vector.

When $\{\varepsilon_t\}_{t\in\mathbb{Z}}$ are *iid* for some $\beta_0 \in \Theta_0$, and independent of the covariates, it is well know that $\left\{\sqrt{T}\hat{\rho}_{T\beta_0}(\tau)\right\}_{\tau=1}^{m}$ are asymptotically independent distributed as standard normal. However, there are other serial dependence cases such that H_0 holds though the sample autocorrelations are not asymptotically *iid*. In fact, independence of the true error $\{\varepsilon_t\}_{t\in\mathbb{Z}}$ with respect to $\{\mathbf{X}_t\}_{t\in\mathbb{Z}}$ is a stronger condition than needed to have the multiplicative error model representation as (1) and we do not carry this assumption through the rest of this paper. Also, higher order serial dependence may be expected, and we do not make assumptions about its possible forms.

Define the vector containing the first m residuals sample autocorrelation

$$\hat{\boldsymbol{\rho}}_{T\beta}^{(m)} = (\hat{\rho}_{T\beta}(1), \ldots, \hat{\rho}_{T\beta}(m))'.$$

Under H_0, but allowing high-order dependence on ε_t,

$$\sqrt{T}\hat{\boldsymbol{\rho}}_{T\beta_0}^{(m)} \xrightarrow{d} N\left(0, \boldsymbol{A}_{\beta_0}^{(m)}\right)$$

where $\boldsymbol{A}_{\beta_0}^{(m)}$ is a $m \times m$ positive definite variance-covariance matrix, see e.g. Romano and Thombs (1996). It is important to emphasize here that we are not imposing any *ad hoc* restrictions on the structure of $\boldsymbol{A}_{\beta_0}^{(m)}$, hence allowing for unknown forms of heteroskedasticity and non-zero cross terms.

Consider the vector of re-scaled sample autocorrelations,

$$\tilde{\boldsymbol{\rho}}_{T\beta}^{(m)} = (\tilde{\rho}_{T\beta}^{(m)}(1), \ldots, \tilde{\rho}_{T\beta}^{(m)}(m))' = \hat{\boldsymbol{A}}_{T\beta_0}^{(m)-1/2} \hat{\boldsymbol{\rho}}_{T\beta}^{(m)},$$

where $\hat{\boldsymbol{A}}_{T\beta_0}^{(m)}$ is a $m \times m$ positive definite matrix of statistics such that $\hat{\boldsymbol{A}}_{T\beta_0}^{(m)} = \boldsymbol{A}_{\beta_0}^{(m)} + o_p(1)$. Thus, under H_0 and some regularity conditions, $\sqrt{T}\tilde{\boldsymbol{\rho}}_{T\beta}^{(m)} \xrightarrow{d} N(0, \boldsymbol{I}_m)$.

In practice, we need a preliminary estimator of $\boldsymbol{\beta}_0$. Assume that an estimator $\hat{\boldsymbol{\beta}}_T$ is available such that, under H_0 of no serial autocorrelation,

$$\hat{\boldsymbol{\beta}}_T = \boldsymbol{\beta}_0 + O_p(T^{-1/2}) \tag{4}$$

and

$$\hat{\boldsymbol{A}}_{T\beta_0}^{(m)-1/2} = \boldsymbol{A}_{\beta_0}^{(m)-1/2} + o_p(1). \tag{5}$$

Preliminary $\sqrt{T}$-consistent estimators of β_0 are available in abundant supply, see e.g. Davis et al. (2000), Davis et al. (2003), Fokianos and Tjøstheim (2011), among others. With respect to condition (5) one can consider the Newey-West type estimator of Lobato et al. (2002), using the multiplicative residuals $\hat{\varepsilon}_t$.

Also assume the following regularity conditions:

Assumption 1 $(Y_t, \mathbf{X}_t', \varepsilon_t)'$ *is strictly stationary,* ε_t *has mean 1,* $E|\varepsilon_t|^{4+2\delta} < \infty$, *for some* $\delta > 0$, *and* $(Y_t, \boldsymbol{X}_t', \varepsilon_t)'$ *is strong mixing with coefficients* α_j *satisfying* $\sum_{j=1}^{\infty} \alpha_j^{\delta/(2+\delta)} < \infty$, *where,*

$$\alpha_j = \sup_{A,B} |\Pr(AB) - P(A)P(B)|$$

and A *and* B *vary over events in the* σ*-fields generated by* $\{(Y_t, X_t', \varepsilon_t)', t \leq 0\}$ *and* $\{(Y_t, X_t', \varepsilon_t)', t \geq j\}$.

Assumption 2 *For all $\boldsymbol{\beta} \in \boldsymbol{\Theta} \subset \mathbb{R}^k, \boldsymbol{\Theta}$ compact, $E||\lambda_t^{-1}\boldsymbol{X}_t'||^{4+\delta} < \infty$, $E||\lambda_t^{-1}\boldsymbol{X}_t\boldsymbol{X}_t'||^{4+\delta} < \infty$, for some $\delta > 0$.*

Next proposition provides an asymptotic expansion for $\sqrt{T}\tilde{\boldsymbol{\rho}}_{T\beta_T}^{(m)}$, which implies that under H_0, conditions (4)-(5), and Assumptions 1-2, $\sqrt{T}\tilde{\boldsymbol{\rho}}_{T\beta_T}^{(m)}$ converges to a vector of independent normal variables plus a stochastic drift, which depends on the estimation effect, $\hat{\boldsymbol{\beta}}_T - \boldsymbol{\beta}_0$. Define

$$\boldsymbol{\xi}_\beta^{(m)} = A_\beta^{(m)-1/2}\boldsymbol{\zeta}_\beta^{(m)}$$

with $\boldsymbol{\xi}_\beta^{(m)} = \left(\boldsymbol{\xi}_\beta(1)', \ldots, \boldsymbol{\xi}_\beta(m)'\right)'$ and $\boldsymbol{\zeta}_\beta^{(m)} = \left(\boldsymbol{\zeta}_\beta(1)', \ldots, \boldsymbol{\zeta}_\beta(m)'\right)'$, such that $\boldsymbol{\zeta}_\beta$ is defined by

$$\frac{\partial}{\partial \boldsymbol{\beta}}\hat{\rho}_{T\beta}(j) \xrightarrow{p} \boldsymbol{\zeta}_\beta(j) \quad \text{each} \quad j \in \mathbb{Z} \setminus \{0\}$$

under H_0.

Proposition 1 *Under H_0 of no autocorrelation, conditions (4) and (5), and Assumptions (1) and (2),*

$$\tilde{\boldsymbol{\rho}}_{T\hat{\beta}_T}^{(m)} = \tilde{\boldsymbol{\rho}}_{T\beta_0}^{(m)} + \boldsymbol{\xi}_{\beta_0}^{(m)}(\hat{\boldsymbol{\beta}}_T - \boldsymbol{\beta}_0) + o_p(T^{-1/2})$$

Proof See Appendix.

From the proof of Proposition 1, one can see that if $\{\mathbf{X}_t\}_{t\in\mathbb{Z}}$ independent of $\{\varepsilon_t - 1\}_{t\in\mathbb{Z}}$, $\boldsymbol{\xi}_{\beta_0}^{(m)}$ would be zero. This is still the case if $\{\mathbf{X}_t\}_{t\in\mathbb{Z}}$ is strictly exogenous. Hence, with this strong assumption, asymptotically there is no effect of using estimated parameters in the residuals sample autocorrelation. This is precisely the case considered by Davis et al. (2000). However, once we relax the strictly exogeneity assumption to the case where $\{\varepsilon_t - 1\}_{t\in\mathbb{Z}}$ is a martingale difference with respect to the σ-field generated by $\{\boldsymbol{X}_s, t \leq s\}$, $\boldsymbol{\xi}_\beta^{(m)}$ would not be zero since $\varepsilon_{t-\tau}$, might be correlated with λ_t and $\mathbf{X}_t$. This also would be the case when $\{\varepsilon_t - 1\}_{t\in\mathbb{Z}}$ is contemporaneously uncorrelated with $\{\mathbf{X}_t\}_{t\in\mathbb{Z}}$. Hence, departures from independence of the errors with respect to the covariates lead to an additional stochastic drift on the estimated residual autocorrelation due to the estimation effect.

The asymptotic distribution of $\sqrt{T}\tilde{\boldsymbol{\rho}}_{T\beta_T}^{(m)}$, under H_0, could be derived from the asymptotic joint distribution of $\left\{\sqrt{T}\tilde{\boldsymbol{\rho}}_{T\beta_0}^{(m)}, \sqrt{T}\left(\hat{\boldsymbol{\beta}}_T - \boldsymbol{\beta}_0\right)\right\}$, under suitable conditions. Nonetheless, different models and estimators would require different derivations, which can be cumbersome.

Instead of adopting this approach, we suggest an asymptotically distribution-free transform of the estimated residuals sample autocorrelation by means of recursive least squares projections, as proposed by Delgado and Velasco (2011).

3. A martingale transform of the residuals sample autocorrelation function with estimated parameters

In order to deal with the distribution of residual autocorrelation with estimated parameters, Delgado and Velasco (2011) propose a transformation based on the recursive least squares residuals introduced by Brown et al. (1975) for CUSUM tests of parameter instability. In order to motivate the transformation, consider the asymptotic decomposition in Proposition 1,

$$\tilde{\rho}^{(m)}_{T\hat{\beta}_T}(\tau) = \tilde{V}^{(m)}_{T\hat{\beta}_T}(\tau) + o_p(T^{-1/2}), \qquad \tau = 1, \ldots, m,$$

with

$$\tilde{V}^{(m)}_{T\hat{\beta}_T}(\tau) = \tilde{\rho}^{(m)}_{T\beta_0}(\tau) + \boldsymbol{\xi}_{\beta_0}(\tau)(\hat{\boldsymbol{\beta}}_T - \boldsymbol{\beta}_0).$$

The source of asymptotic autocorrelation, under H_0, in $\left\{\tilde{V}^{(m)}_{T\hat{\beta}_T}(\tau)\right\}_{\tau=1}^{m}$ is $(\hat{\boldsymbol{\beta}}_T - \boldsymbol{\beta}_0)$. Then, the transformation consists in using a linear operator $\mathcal{L}^{(m)}$ such that $\left\{\mathcal{L}^{(m)}\tilde{V}^{(m)}_{T\hat{\beta}_T}\right\}_{\tau\geq 1}$ are asymptotically uncorrelated. Delgado and Velasco (2011) considered the operator that transform any sequence $\{\eta(\tau)\}_{\tau=1}^{m}$ in the forward recursive residuals of its least square projection on $\left\{\boldsymbol{\xi}_{\beta_0}(\tau)\right\}_{\tau=1}^{m}$,

$$\mathcal{L}^{(m)}\eta(\tau) = \eta(\tau) - \boldsymbol{\xi}_{\beta_0}(\tau)\left(\sum_{l=\tau+1}^{m} \boldsymbol{\xi}_{\beta_0}(l)'\boldsymbol{\xi}_{\beta_0}(l)\right)^{-1} \sum_{l=\tau+1}^{m} \boldsymbol{\xi}_{\beta_0}(l)'\eta(l).$$

Backward recursive residuals could also be alternatively used.

Notice than, when it is applied to $\left\{\tilde{V}^{(m)}_{T\hat{\beta}_T}(\tau)\right\}_{\tau=1}^{m}$, we have $\mathcal{L}^{(m)}\tilde{V}^{(m)}_{T\hat{\beta}_T}(\tau) = \mathcal{L}^{(m)}\tilde{\rho}^{(m)}_{T\beta_0}(\tau)$, $\tau = 1, \ldots m - k$, which does not depend on $(\hat{\boldsymbol{\beta}}_T - \boldsymbol{\beta}_0)$. Since $\left\{\sqrt{T}\tilde{\rho}^{(m)}_{T\beta_0}(\tau)\right\}_{\tau\geq 1}$ are asymptotically distributed as *iid* standard normal, $\left\{\sqrt{T}\mathcal{L}^{(m)}\tilde{\rho}^{(m)}_{T\beta_0}(\tau)\right\}_{\tau\geq 1}$ are asymptotically distributed as independent normal random variables with mean zero and variance

$$\sigma_\rho^2(\tau) = 1 + \boldsymbol{\xi}_{\beta_0}(\tau)\left(\sum_{l=\tau+1}^{m} \boldsymbol{\xi}_{\beta_0}(l)'\boldsymbol{\xi}_{\beta_0}(l)\right)^{-1} \boldsymbol{\xi}_{\beta_0}(\tau)'. \tag{6}$$

In practice, we need a consistent estimator of $\boldsymbol{\xi}_{\beta_0}$ to perform the transformation. As it is shown in the proof of Proposition 1, under H_0 we have

$$\left\| \frac{\partial}{\partial \boldsymbol{\beta}} \hat{\boldsymbol{\rho}}_{T\beta_0}(\tau) - \frac{1}{\boldsymbol{\gamma}_{\beta_0\varepsilon}(0)} \frac{\partial}{\partial \boldsymbol{\beta}} \hat{\boldsymbol{\gamma}}_{T\beta_0}(\tau) \right\| = o_p(1), \ \tau \neq 0.$$

Thus, standardizing by $\hat{\gamma}_{T\beta_0\varepsilon}(0)$ in $\hat{\boldsymbol{\rho}}_{T\beta_0}$ has no asymptotic effect on $\boldsymbol{\zeta}_\beta$. Then, we can estimate $\boldsymbol{\xi}_{\beta_0}^{(m)}$ by

$$\hat{\boldsymbol{\xi}}_{T\hat{\beta}_T}^{(m)} = \hat{A}_{T\hat{\beta}_T}^{(m)-1/2} \hat{\boldsymbol{\zeta}}_{T\hat{\beta}_T}^{(m)}, \tag{7}$$

where $\hat{\boldsymbol{\xi}}_{T\hat{\beta}_T}^{(m)} = \left(\hat{\boldsymbol{\xi}}_{T\hat{\beta}_T}(1)', \ldots, \hat{\boldsymbol{\xi}}_{T\hat{\beta}_T}(m)'\right)'$ and $\hat{\boldsymbol{\zeta}}_{T\hat{\beta}_T}^{(m)} = \left(\hat{\boldsymbol{\zeta}}_{T\hat{\beta}_T}(1)', \ldots, \hat{\boldsymbol{\zeta}}_{T\hat{\beta}_T}(m)'\right)'$, with

$$\hat{\boldsymbol{\zeta}}_{T\hat{\beta}_T}(\tau) = \frac{1}{T\ \hat{\gamma}_{T\hat{\beta}_T}(0)} \left(\sum_{t=\tau+1}^{T} \frac{Y_{t-\tau}\left(Y_t - \lambda_\tau\right)\boldsymbol{X}_{t-\tau}' - Y_t\left(Y_{t-\tau} - \lambda_{\tau-\tau}\right)\boldsymbol{X}_t'}{\lambda_t \lambda_{t-\tau}} \right)$$

$$\hat{\gamma}_{T\hat{\beta}_T}(\tau) = \frac{1}{T} \sum_{t=\tau+1}^{T} \frac{\left(Y_t - \hat{\lambda}_t\right)\left(Y_{t-\tau} - \hat{\lambda}_{t-\tau}\right)}{\hat{\lambda}_t \hat{\lambda}_{t-j}}.$$

The feasible transformation consists of the operator $\hat{\mathcal{L}}_T^{(m)}$, which transforms any sequence $\{\eta(\tau)\}_{\tau=1}^m$ in the forward recursive residuals of its least square projection on $\left\{\hat{\boldsymbol{\xi}}_{T\hat{\beta}_T}^{(m)}\right\}_{\tau=1}^m$,

$$\hat{\mathcal{L}}_T^{(m)}\eta(\tau) = \eta(\tau) - \hat{\boldsymbol{\xi}}_{T\hat{\beta}_T}(\tau)\left(\sum_{l=\tau+1}^{m} \hat{\boldsymbol{\xi}}_{T\hat{\beta}_T}(l)'\hat{\boldsymbol{\xi}}_{T\hat{\beta}_T}(l)\right)^{-1} \sum_{l=\tau+1}^{m} \hat{\boldsymbol{\xi}}_{T\hat{\beta}_T}(l)'\eta(l).$$

The transformed residuals sample autocorrelations, in the presence of estimated parameters is

$$\bar{\rho}_{T\beta}^{(m)}(\tau) = \frac{\hat{\mathcal{L}}_T^{(m)} \tilde{\rho}_{T\beta}^{(m)}(\tau)}{\hat{\sigma}_{T\rho}(\tau)}, \tau = 1\ldots, m-k, \tag{8}$$

where $\hat{\sigma}_{T\rho}^2(\tau) = 1 + \hat{\boldsymbol{\xi}}_{T\hat{\beta}_T}(\tau)\left(\sum_{l=\tau+1}^{m} \hat{\boldsymbol{\xi}}_{T\hat{\beta}_T}(l)'\hat{\boldsymbol{\xi}}_{T\hat{\beta}_T}(l)\right)^{-1} \hat{\boldsymbol{\xi}}_{T\hat{\beta}_T}(\tau)'$ is the estimator of $\sigma_\rho^2(\tau)$. Notice than we can only transform the first $m-k$ sample autocorrelations, because, giving a scaling matrix $\hat{\boldsymbol{A}}^{(m)}_{T\hat{\beta}_T}$, there are no more degrees of freedom available when k parameters are estimated.

As discussed by Delgado and Velasco (2011), we could also use backward residuals, but with this approach, we would lose the first k residuals sample autocorrelations, which usually are the most informative.

In order to prove that, under H_0, $\bar{\boldsymbol{\rho}}^{(m)}_{T\hat{\beta}_T} = \left(\bar{\rho}^{(m)}_{T\hat{\beta}_T}(1), \ldots, \bar{\rho}^{(m)}_{T\hat{\beta}_T}(m-k)\right)'$ and $\bar{\boldsymbol{\rho}}^{(m)}_{T\beta_0}$ are asymptotically equivalent, and $\sqrt{T}\bar{\boldsymbol{\rho}}^{(m)}_{T\beta_0}$ is asymptotically distributed as a vector of independent standard normals, we need an extra technical assumption in order to compute the transform.

Assumption 3 *For $m > k$,*

$$\sum_{l=1+m-k}^{m} \boldsymbol{\xi}_{T\hat{\beta}_T}(l)' \boldsymbol{\xi}_{T\hat{\beta}_T}(l)$$

is positive definite.

Theorem 1 *Under H_0, $m > k$, Assumptions 1-3, and with $\hat{\boldsymbol{\beta}}_T$ satisfying (4) and (5),*

$$\bar{\boldsymbol{\rho}}^{(m)}_{T\hat{\beta}_T} = \bar{\boldsymbol{\rho}}^{(m)}_{T\beta_0} + o_p(T^{-1/2})$$

and

$$\sqrt{T}\bar{\boldsymbol{\rho}}^{(m)}_{T\beta_0} \xrightarrow{d} N_{m-k}(0, \boldsymbol{I}_{m-k}).$$

Proof See Appendix.

Theorem 1 forms the basis for implementing asymptotic test of lack of autocorrelation based on the asymptotically *iid* sequence $\bar{\boldsymbol{\rho}}^{(m)}_{T\hat{\beta}_T}$, as described in the next section.

4. Testing lack of autocorrelation on the multiplicative residuals with estimated parameters

Our goal is to test if there is evidence that the multiplicative error is not autocorrelated, which provides a model check. That is, we seek to test

$$H_0 : \rho_{\beta_0}(\tau) = 0,\ \tau = 1, \ldots, m \text{ and some } \beta_0 \in \Theta \tag{9}$$

against the fixed alternative hypothesis

$$H_1^s = \rho_{\beta_0}(\tau) \neq 0 \text{ for some } \tau = 1, \ldots, m \text{ for any } \beta_0 \in \Theta \tag{10}$$

for some $m > k$.

One of the most popular test statistic for lack of autocorrelation as expressed by H_0 is the Portmanteau Box and Pierce (1970) statistic, $\hat{B}^{BP}_{T\hat{\beta}_T}(s)$, with

$$\hat{B}^{BP}_{T\beta}(s) = T \sum_{\tau=1}^{s} \hat{\rho}_{T\beta}(\tau)^2,$$

for some $k < s \leq m$, where k is the number of estimated parameters in the model.

Ljung and Box (1978) propose a small modification to $\hat{B}^{BP}_{T\hat{\beta}_T}(s)$ in order to have better finite sample properties. The test statistic proposed by Ljung and Box (1978) is

$$\hat{B}^{LB}_{T\hat{\beta}_T}(s) = T(T+2) \sum_{\tau=1}^{s} \frac{\hat{\rho}_{T\beta}(\tau)^2}{T-\tau}.$$

The Portmanteau type tests of Box and Pierce (1970) assumes that s is a fixed number. This restriction leads to tests which are not able to detect serial correlation appearing at lags larger than s. In order to overcome this issue, Hong (1996) allows s growing with the sample size. Nonetheless, although consistent, omnibus autocorrelation tests present low empirical power when the autocorrelation appears at higher lags, though large s, typically of $O\left(T^{-1/2}\right)$, is needed in order to get reasonable size accuracy - see Escanciano (2009) and references therein for the theoretical explanations.

On the other hand, the Portmanteau tests of Box and Pierce (1970) fall into the class of Neyman's smooth test, which are optimal to detect fixed local alternatives of the type

$$H_{1T:}\ \rho_{\beta_0}(\tau) = \frac{r(j)}{\sqrt{T}} + \frac{j_T(\tau)}{T} \forall j = 1, 2, \ldots, m \tag{11}$$

where r and j_T are square summable such that ρ_{β_0} is positive definite sequence for all T - see Delgado and Velasco (2010). If one has a particular local alternative in mind, Delgado and

Velasco (2010) propose a class of specification tests which maximize the power function when testing in the direction of the chosen local alternative. Nonetheless, Escanciano and Lobato (2009) show that, when one does not have an alternative r in mind, the Neyman's smooth tests, and hence the BP type test, are optimally adaptive to the unknown local alternative.

We consider a test of H_0 based on the sums of squared transformed autocorrelations, which is a version of the BP test statistic replacing the sample residual autocorrelation by its asymptotically distribution-free transformation in (8)

$$\bar{B}_{T\beta}^{(m)}(s) = T \sum_{\tau=1}^{s} \bar{\rho}_{T\beta}^{(m)}(\tau)^2,$$

for some $1 \leq s \leq m - k$.

It follows from Theorem 1 that, under H_0,

$$\bar{B}_{T\hat{\beta}_T}^{(m)}(s) \xrightarrow{d} \sum_{\tau=1}^{s} Z_\tau^2 \equiv \chi_{(s)}^2$$

where $\{Z_\tau\}_{\tau\in\mathbb{N}}$ are *iid* standard normals, and $\chi_{(s)}^2$ is a chi-square distribution with s degrees of freedom. This result is summarized in the Corollary 1

Corollary 1 *Under H_0 and the conditions stated on Theorem 1,*

$$\bar{B}_{T\hat{\beta}_T}^{(m)}(s) \xrightarrow{d} \boldsymbol{\chi}_{(s)}^2,$$

$1 \leq s \leq m - k$.

Box and Pierce (1970) show that, when $\{\varepsilon_t\}_{t\in\mathbb{Z}}$ are *iid*, $s > k$, and s is increasing with T, $\hat{B}_{T\beta}^{BP}(s) \overset{asy}{\sim} \chi_{(s-k)}^2$. However, when s remains fixed, $\hat{B}_{T\beta}^{BP}(s)$ has a limiting distribution that depends on unknown features of the data generating process such as the parameter vector $\boldsymbol{\beta}_0$, preventing its use when s is small. In fact, the residual sample autocorrelation at any lag is an inconsistent estimator of the true innovation autocorrelation function.

When $\{\varepsilon_t\}_{t\in\mathbb{Z}}$ exhibits high-order dependence, the asymptotic variance of the residuals sample autocorrelations is cumbersome to calculate, as it is shown by Romano and Thombs

(1996) in a weak ARMA model context. In this case, Box and Pierce (1970) test statistic is no long well approximated by a $\chi^2_{(s-k)}$ distribution since $\rho^{(m)}_{T\beta}(\tau)$ is no longer asymptotically distributed as a standard normal random variable.

On the other hand, our proposed test statistic $\bar{B}^{(m)}_{T\beta}(s)$ prevents for $\{\varepsilon_t\}_{t\in\mathbb{Z}}$ exhibiting high-order dependence. Also, contrary to Box and Pierce (1970), the test statistic $\bar{B}^{(m)}_{T\beta}(s)$ is pivotal for fixed s, since the estimated parameters effect is already projected out. Hence, even when s is fixed, and $\{\varepsilon_t\}_{t\in\mathbb{Z}}$ exhibits high-order dependence, our proposed test statistic follows a $\boldsymbol{\chi}^2_{(s)}$ asymptotically, $m \geq s+k$.

In order to discuss the power of the proposed test, consider the local alternatives of the form (11), where r and j_T are square summable such that ρ_{β_0} is positive definite sequence for all T.

Concerning the asymptotic distribution of $\bar{\boldsymbol{\rho}}^{(m)}_{T\beta}$ under H_{1T}, define the vector of projected and standardized autocorrelation drifts as $\check{\boldsymbol{h}}^{(m)}_{\beta} = \left(\check{h}^{(m)}_{\beta}(1), \ \ldots, \ \check{h}^{(m)}_{\beta}(m-k)\right)'$, where

$$\check{h}^{(m)}_{\beta}(\tau) = h^{(m)}_{\beta}(\tau) - \boldsymbol{\xi}_{\beta}(\tau)\left(\sum_{l=\tau+1}^{m} \boldsymbol{\xi}_{\beta}(l)'\boldsymbol{\xi}_{\beta}(l)\right)^{-1} \sum_{l=\tau+1}^{m} \boldsymbol{\xi}_{\beta}(l)' h^{(m)}_{\beta}(\tau) \tag{12}$$

$\tau = 1, \ldots, m-k$,

$$h^{(m)}_{\beta}(\tau) = \sum_{i=1}^{m} \left[A^{(m)-1/2}_{\beta}\right]_{(\tau,i)} r(i). \tag{13}$$

and let $\bar{\boldsymbol{h}}^{(m)}_{\beta} = \left(\bar{h}^{(m)}_{\beta}(1), \ \ldots, \ \bar{h}^{(m)}_{\beta}(m-k)\right)'$,where $\bar{h}^{(m)}_{\beta}(\tau) = \check{h}^{(m)}_{\beta}(\tau)/\sigma_{\rho}(\tau)$, with $\sigma_{\rho}(\tau)$ as defined in (6).

Theorem 2 *Under H_{1T}, $m > k$, Assumptions 1-3, and with $\hat{\boldsymbol{\beta}}_T$ satisfying (4) and (5),*

$$\bar{\boldsymbol{\rho}}^{(m)}_{T\hat{\beta}_T} = \bar{\boldsymbol{\rho}}^{(m)}_{T\beta_0} + o_p(T^{-1/2})$$

and

$$\sqrt{T}\bar{\boldsymbol{\rho}}^{(m)}_{T\beta_0} \xrightarrow{d} N_{m-k}(\bar{\boldsymbol{h}}^{(m)}_{\beta_0}, I_{m-k}).$$

Proof See Appendix.

From Theorem 2, we can see that the sample transforms of the residuals sample autocorrelations are asymptotically distributed as Normal with non-zero mean. Then, it follows that, under nonparametric local alternatives of the form (11), our proposed test statistic $\bar{B}^{(m)}_{T\hat{\beta}_T}(s)$ is asymptotically distributed as non-central chi-squared, $\chi^2_{(s)}(\varphi)$, with noncentrality parameter φ equal to $\sum_{j=i}^{s}\left(\bar{h}^{(m)}_{\beta_0}(j)\right)^2$. Hence, our test is able to detect nonparametric local alternatives like H_{1T}, which converges to the null hypothesis at the parametric rate. The classical BP test, as shown by Hong (1996), does not meet this property. Moreover, our test is consistent against fixed alternatives of the form (10). We summarize these results in Corollary 2.

Corollary 2 *Under H_1 and the conditions stated on Theorem 2,*

$$\bar{B}^{(m)}_{T\hat{\beta}_T}(s) \xrightarrow{d} \boldsymbol{\chi}^2_{(s)}(\varphi),$$

$1 \leq s \leq m-k$, $\varphi = \sum_{j=i}^{s}\left(\bar{h}^{(m)}_{\beta_0}(j)\right)^2$. *Moreover, under fixed alternatives of the form (10), for all $c < \infty$,*

$$\lim_{T\to\infty} P\left[\bar{B}^{(m)}_{T\hat{\beta}_T}(s) > c\right] = 1$$

5. Monte Carlo Simulations

This section illustrates the finite sample performance of our proposal comparing the simulated empirical percentage of rejections under H_0 and H_1 of alternative residual sample autocorrelations based tests. We consider sample sizes $T = 100$ and 300, and $10,000$ replications in each experiment. All models are estimated using a Poisson Quasi-Likelihood.

For $t = 1, \ldots, T$, we consider the following null models:

$$Y_t \sim Poisson(\lambda_t)$$

and

$$\lambda_t = \exp\left(1 + X_t + rv_t\right),$$

$$X_t = 0.5X_{t-1} + u_t,$$

and $\{u_t\}_{t\in\mathbb{Z}}$ follows an *iid* standard normal distribution, and $\{v_t\}_{t\in\mathbb{Z}}$ follows and *iid* normal distribution with mean -0.347 and variance 0.693. Hence, $\exp(v_t)$ follows a log-normal distribution with mean 1 and variance 1. We consider two specifications: (a) $r = 0$, and (b) $r = 1$. This way, on specification (a) we have a standard Poisson model, and on the (b) we introduce a multiplicative latent process which leads to overdispersion, as first considered by Zeger (1988).

In both specifications, the conditional mean of the count process is given by λ_t, and hence under H_0 there is a centered multiplicative error $\varepsilon_t - 1$ with mean 0. Notice that (a) and (b) leads to different conditional heteroskedasticity forms: residual conditional variance on specifications (a) is equal to λ_t^{-1}, and on (b) is equal to $1 + \lambda_t^{-1}$.

Our test statistic $\bar{B}^{(m)}_{T\hat{\beta}_T}(s)$ uses critical values from a chi squared distribution with s degrees of freedom. The nominal level of all tests is 5%. For the sake of comparison, we use values of s from 1 up to 15 and 20 for $T = 100$ and 300 respectively. We set $m = s + k$ in order to avoid studentization on unnecessary residual autocorrelation. Moreover, $A^{(m)}_{\beta_0}$ need to be estimated and hence it is not reasonable to set larger m than needed. We use three estimates of $A^{(m)}_{\beta_0}$: (i) $\hat{\boldsymbol{A}}^{(m)}_{T\hat{\beta}_T} = I_m$, (ii) $\hat{\boldsymbol{A}}^{(m)}_{T\hat{\beta}_T} = diag\left\{\hat{a}^{(1,1)}_{T\hat{\beta}_T}, \ldots, \hat{a}^{(m,m)}_{T\hat{\beta}_T}\right\} / \hat{\gamma}_{T\hat{\beta}_T\varepsilon}(0)^2$, with $\hat{a}^{(j,j)}_{T\beta} = T^{-1}\sum_{t=1+j}^{T} \varepsilon^2_{t\beta}\varepsilon^2_{t-j\beta}$, and (iii) the Newey-West-type unrestricted estimator of $A^{(m)}_{\beta_0}$ used by Lobato et al. (2002) with preliminary bandwidth $n = 2\left(T/100\right)^{1/3}$, no prewhitening and Barlett's kernel, as in Delgado and Velasco (2011).

We compare our new test statistic with the classical BP test, $\hat{B}^{BP}_{T\hat{\beta}_T}(s)$, under *iid* multiplicative innovations (with $\chi^2_{(s-k)}$ approximation), and with a BP test variation, $\hat{B}^{BP\ D}_{T\hat{\beta}_T}(s)$, where we standardize the residuals sample autocorrelations by $\hat{\boldsymbol{A}}^{(m)}_{T\hat{\beta}_T}$, where $\hat{\boldsymbol{A}}^{(m)}_{T\hat{\beta}_T}$ is considered to be diagonal as in (ii).

Figure 1 report the simulated empirical size of the tests. We can observe that, for both versions of the classical Box-Pierce test, the type I error is out of control for any sample size. In general, in both specifications and sample sizes, tests based on recursive projections control the type I error, with the exception on specification (b), when $T = 100$, high values of s are used with unrestricted estimator of $\boldsymbol{A}^{(m)}_{T\hat{\beta}_T}$, perhaps due to the need of inverting a matrix of larger dimension. Nonetheless, when $T = 300$, this over-sizing distortions disappears.

$\boxed{\text{INSERT FIGURE 1 HERE}}$

We consider also two specifications under the alternative, using the GARMA(0,1) of Davis et al. (2003) and Benjamin et al. (2003), and introducing an autocorrelated latent process on the Poisson parameter, as in Zeger (1988). More precisely, we consider the following specifications under H_1:

$$
\begin{aligned}
&(c)\ Y_t \sim Poisson(\omega_t),\ \omega_t = \exp\left(1 + X_t + 0.5\frac{Y_{t-1} - \omega_{t-1}}{\omega_{t-1}}\right) \\
&(d)\ Y_t \sim Poisson(\pi_t),\ \pi_t = \exp\left(1 + X_t + z_t\right))
\end{aligned}
$$

where

$$z_t = 0.62 z_{t-1} + v_t.$$

Since $\{\exp(z_t)\}$ is log-normal, specification (d) leads to first order residual autocorrelation of 0.5.

$\boxed{\text{INSERT FIGURE 2 HERE}}$

The two considered versions of the classical Box-Pierce test do not control size, and hence we only report the rejections under the alternative using our test statistic $\bar{B}^{(m)}_{T\hat{\beta}_T}(s)$. We can see in Figure 2 that our test $\bar{B}^{(m)}_{T\hat{\beta}_T}(s)$ based on recursive residuals exhibit good power performance for all s considered, for both the $\hat{\boldsymbol{A}}^{(m)}_{T\hat{\beta}_T}$ being diagonal or the identity matrix. When using the the robust estimator of $\boldsymbol{A}^{(m)}_{T\hat{\beta}_T}$, we can see. as expected, that our test statistic loses power when considering high values of s.

6. Risk Management and U.S. Corporate Bankruptcies

In order to illustrate the appealing of our proposed test statistic in applied econometrics, we analyze different specifications of common credit risk models.

In a seminal paper, Das et al. (2007) analyze if the observable variables are sufficient to explain the default time correlation of U.S. non-financial corporations. Using a test statistic based on the count of defaults in a given period, Das et al. (2007) reject the hypothesis of defaults being conditional independent, suggesting some evidence of excess default clustering. This finding has important implications for practitioners because many popular default risk models rely on the assumption of conditionally independent defaults. Moreover, as shown by Duffie et al. (2009), ignoring such default clustering leads to substantial downward bias on extreme default losses probabilities.

In order to overcome such consequences, Duffie et al. (2009) propose to add a common dynamic "frailty" effect on the firms default hazard, that is, an unobserved correlated latent process common to all firms. As an alternative to the duration model of Duffie et al. (2009), Koopman et al. (2011, 2012), using time series count data panel models, propose new estimators for the measurement and forecasting of default probabilities when excess default clustering is present, allowing for a large number of macroeconomic and financial variables, an industry fixed effects and a common frailty effect. Differently than Duffie et al. (2009), which model the frailty effect as continuous-time process, Koopman et al. (2011, 2012) rely on a state space specification, such that the frailty effect is modeled as a Gaussian AR(1). Their results confirms the findings of Das et al. (2007) in the sense that there is some evidence of a correlated frailty effect. However, an important question remain unanswered: Is the AR(1) latent process structure enough to capture all the excess default? In other words, is there any evidence of residuals serial correlation, after including this additional parameter?

Our test for lack of autocorrelation is a valuable tool in order to access if the proposed model for bankruptcy counts is correctly specified. Within our approach, we are able to test both if there is evidence of excess correlation, and, in case there is, if the usual assumption that considering only first order dynamics is enough to capture the excess of default/bankruptcy correlation. This second hypothesis, to the best of our knowledge, has not been verified so far. This is an important model check since, as pointed out by Koopman et al. (2011), model misspecification can lead to underestimation of corporate risk.

When the interest is on determining adequate economic capital buffers, the focus of the analysis is on aggregate default or bankruptcies rather than on firm specific default. A modeling strategy that deals directly with aggregate default counts is a natural alternative from the procedure of Duffie et al. (2009) and Koopman et al. (2011, 2012), in which they first estimate the firms default probability and then aggregate.

With this in mind, using monthly data on bankruptcy filed in the United States Bankruptcy Courts from January 1985 until October 2012, available from UCLA-LoPucki Bankruptcy Research Database [1], we analyze if there is any evidence of excess correlation in bankruptcy counts. Moreover, we test for lack of autocorrelation in the residuals of different Poisson-GARMA models, particularly if including an AR(1) or MA(1) term is enough to capture the linear dynamics of the monthly bankruptcy counts. Although bankruptcy data is available since October 1979, we only use data from 1985 onwards, that is, only the period after the "Great Moderation". We do it in order to avoid the presence of well documented structural breaks.

UCLA-LoPucki Bankruptcy Research Database contains data on all large, public company bankruptcy cases filed in the United States Bankruptcy Courts. By large firms, they consider firms which have declared assets of more than \$100 million, measured in 1980 dollars, the year before the firm filled the bankruptcy case. A firm is considered public if they report to the Securities Exchange Commission in the last three years prior to bankruptcy. Following Compustat, although only 22% of the public firms has higher market value than \$100 Million in 2011, these firms represent 70% of total assets and sales of all firms listed, and hence represent an important category of firms. Monthly bankruptcy counts are considered in terms of the month the bankruptcy file was filed.

Macroeconomic and financial monthly data are obtained from the St. Louis Fed online database FRED, see Table 1 for a listing of macroeconomic and financial data. These involve business cycle measurements, labor market conditions, interest rate and credit spread and are typically used in macro stress test - see Tarullo (2010) for instance. Variables are expressed as yearly growth rates (INDPRO, PERMIT, PPIFGS and PPIENG) or as yearly differences (UNRATE, BAA, FEDFUNDS, GS10, SP500RET and SP500VOL). We also consider a time dummy which takes value one after September 2005, in order to capture the

1. Available at http://lopucki.law.ucla.edu/

effect of a major bankruptcy law reform, the Bankruptcy Abuse Prevention and Consumer Protection Act (BAPCPA), signed in October 2005.

INSERT TABLE 1 HERE

We estimate the count data model using the Poisson quasi-likelihood only with the covariates, and also consider the observation-driven Poisson GARMA Model of Davis et al. (2003) with and additional AR(1) or MA(1) parameter - GARMA(1,0) and GARMA(0,1), respectively. We finally provide AIC and BIC values for the models considered. For checking the fit of the models, we use the Box-Pierce test based on the transformed multiplicative residuals autocorrelation, $\bar{B}^{(m)}_{T\hat{\beta}_T}(s)$, with s varying from 1 to 6. These choices include all the usual lag choices in similar applications supported by our simulations, given that $T = 324$. We report the analysis with $\hat{A}^{(m)}_{\hat{\beta}_T}$ being the identity matrix and diagonal. Since it is clear from the simulations that the classical Box-Pierce test, $\bar{B}^{BP}_{T\hat{\beta}_T}(s)$, does not control size, we omit it. The results of the test statistic are presented on Table 2. Estimated parameters for the different models are reported on Table 3.

From the specification tests presented on Table 2 one can conclude that the simple static Poisson model, which contains only the macroeconomic covariates, is strongly rejected by the recursive Portmanteau test $\bar{B}^{(m)}_{T\hat{\beta}_T}(s)$. This result points to the same direction of the results of Duffie et al. (2009) and Koopman et al. (2011, 2012): there is evidence of a bankruptcy cluster beyond the one implied by the macroeconomic variables. As pointed out by these authors, ignoring such excess autocorrelation can lead to mismeasures on risk management, specially underestimation of extreme loss given default.

INSERT TABLE 2 HERE

In order to understand better the source of the rejection of the null for the Poisson model with only covariates, we consider the analysis of individual projected residuals autocorrelations for lags up to 20, with $\hat{\boldsymbol{A}}^{(m)}_{T\hat{\beta}_T}$ being diagonal, and $m = 32$. Recall that transformed autocorrelations can be correctly compared with the usual $\pm 2/\sqrt{T}$ confidence band, as when working with raw data. In Figure 3, we have plotted the autocorrelograms of the transformed residuals of the Poisson model only with covariates. In this plot we can easily identify the source of the rejection, since transformed autocorrelations provide evidence on serial correlation of the underlying innovation from the very first lag onwards.

INSERT TABLE 3 HERE

Once we consider both Poisson GARMA models, we fail to reject H_0, that is, the data supports that these specifications seems to capture the linear dynamics of the bankruptcy counts. These results provides some evidence that, within the exponential functional form, considering just first order autocorrelation is enough to capture the bankruptcy dynamics, as suggested by Duffie et al. (2009), Koopman et al. (2011, 2012) in different contexts. To the best of our knowledge, we are the first to formally test these suggestions.

These results have important implications for risk management. Almost all industry credit risk models, such as CreditMetrics, Moody's KMV and CreditRisk+ rely on the assumption that default and bankruptcies are time independent. However, from the results of our specification testing, we conclude that there is evidence of an excess bankruptcy clustering. The presence of residual autocorrelation may increase bankruptcy rate volatility, and as result it may shift probability mass of an portfolio credit loss distribution toward more extreme values. This would increase capital buffers prescribed by the risk models. Hence, if one ignores the presence of a frailty effect, portfolio credit risk models will tend to be wrong. On the other hand, if one consider first order autocorrelations, as in the GARMA models we have presented, it seems that there is no evidence of model misspecification. This way, we argue that these type of models are more appropriate to model bankruptcies and adjusting the credit risk models for it would not only be relevant for internal risk assessment, but also for external supervision of financial institutions.

7. Conclusion

In this paper, we have proposed a new distribution-free test for lack of autocorrelation in count data models in the presence of estimated parameters, under weak assumptions on the relationship between the covariates and the multiplicative innovations. The test statistic proposed is of the Box and Pierce (1970) type, but contrary to the classical tests, it is able to detect local alternatives converging to the null at the parametric rate. Our test present satisfactory finite sample properties as demonstrated via Monte Carlo simulations. Once our proposal is applied to bankruptcy count models, we rejected the specification of a model with only macroeconomic covariates, but do not reject the null of lack of autocorrelation once we consider dynamic count models as the GARMA(0,1) and the GARMA(1,0). Hence, we advocate that considering this broader class of models seem more appropriate when dealing with bankruptcy risk.

Our basic results can be extended to other situations of practical interest without any additional difficulty. For instance, under suitable conditions, one could consider the multiplicative error model $Y_t = g(\mathcal{F}_{t-1}; \boldsymbol{\theta})\varepsilon_t$ where $g_t(\cdot)$ a known twice differentiable function, $\mathcal{F}_{t-1}$ is the available information at time t (can include lags of Y_t, ε_t, and also a set of covariates $\boldsymbol{X}_t$), $\boldsymbol{\theta}$ is a vector of parameters to be estimated and ε_t has a non-negative distribution with $E\left(\varepsilon_t | \mathcal{F}_{t-1}\right) = 1$. Once $\boldsymbol{\theta}$ is estimated, we can obtain the centered residuals $\left[Y_t - g(\mathcal{F}_{t-1}; \hat{\boldsymbol{\theta}})\right] / g(\mathcal{F}_{t-1}; \hat{\boldsymbol{\theta}}) = \hat{\varepsilon}_t - 1$, and then apply the asymptotically distribution-free transform to the residual sample autocorrelation, and all our results would follow once we properly compute the score of the residual sample autocorrelation.

Regarding the choice of the number of lags included in the test statistic, one can adopt a data-driven procedure based on an AIC/BIC criterion in the lines of Escanciano and Lobato (2009) and Escanciano et al. (2013), at the cost of not being able to detect the kind of local alternatives considered here. The proofs of Escanciano and Lobato (2009) and Escanciano et al. (2013) could be extended to the present case without additional effort.

Acknowledgment

I am most thankful to Miguel Delgado, Carlos Velasco, Peter Robinson, Jesus Gonzalo, Claudio Shikida, Márcio Laurini, Igor Kheiefts and Konstantinos Fokianos for their constructive comments on a previous version of this article. I also would like to thank William Dunsmuir for providing the R code to estimate the Poisson GARMA models. This research was funded by Spanish Plan Nacional de I+D+I grant number ECO2012-3313.

Appendix: Proofs

PROOF OF PROPOSITION 1

It follows from Taylor expansion around the true $\boldsymbol{\beta}$, element by element. For each $j = 1, \ldots, m$, we write

$$\hat{\rho}_{T\hat{\beta}_T}(j) - \hat{\rho}_{T\beta_0}(j) = \frac{\partial \hat{\rho}_{T\beta}(j)}{\partial \boldsymbol{\beta}}(\hat{\boldsymbol{\beta}}_T - \boldsymbol{\beta}_0) + D_T(j)$$

where $D_T(j)$ is

$$D_T(j) = \left(\hat{\boldsymbol{\beta}}_T - \boldsymbol{\beta}_0\right)' \ddot{\boldsymbol{\rho}}_{T\beta^*_{T,j}}(j) \left(\hat{\boldsymbol{\beta}}_T - \boldsymbol{\beta}_0\right),$$

$\ddot{\boldsymbol{\rho}}_{T\beta^*_{T,j}}(j) = \partial^2 \hat{\rho}_{T\beta}(j)/\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}'$ and $\boldsymbol{\beta}^*_{T,j}$ are such that $||\boldsymbol{\beta}^*_{T,j} - \boldsymbol{\beta}_0|| \leq ||\hat{\boldsymbol{\beta}}_T - \boldsymbol{\beta}_0||$, $\forall j = 1, \ldots, m$.

Then, for each $j = 1, \ldots, m$,

$$\frac{\partial}{\partial \beta}\hat{\rho}_{T\beta}(j) = \frac{\frac{\partial}{\partial \beta}\hat{\gamma}_{T\beta}(j)}{\hat{\gamma}_{T\beta}(0)} - \frac{\hat{\gamma}_{T\beta}(j)}{\hat{\gamma}_{T\beta}(0)}\frac{\frac{\partial}{\partial \beta}\hat{\gamma}_{T\beta}(0)}{\hat{\gamma}_{T\beta}(0)}.$$

Using that $\hat{\gamma}_{T\beta_0}(j) = \gamma_{\beta_0}(j) + o_p(1)$, in particular $\gamma_{\beta_0}(j) = 0$ for $j \neq 0$ under H_0 and that

$$\frac{\partial}{\partial \beta}\gamma_{T\beta_0}(0) = -\frac{2}{T}\sum_{t=1}^{T}\left(\frac{Y_t - \lambda_t}{\lambda_t^2}\right)\mathbf{X}_t'$$

$$\xrightarrow{p} -2E\left[\frac{\varepsilon_t - 1}{\lambda_t}\mathbf{X}_t'\right]$$

under Assumptions (1) and (2) and Law of Large Numbers, we conclude that the normalization of $\hat{\boldsymbol{\rho}}^{(m)}_{T\hat{\beta}_T}$ has no asymptotic effect under H_0, so that

$$\frac{\partial}{\partial\boldsymbol{\beta}}\hat{\boldsymbol{\rho}}_{T\beta_0}(j) = \frac{\frac{\partial}{\partial\boldsymbol{\beta}}\gamma_{T\beta_0}(j)}{\gamma_{T\beta_0}(0)} + o_p(1).$$

Without loss of generality, assume that $\gamma_{\beta_0}(0) = 1$. Writing now

$$\begin{aligned}\frac{\partial\gamma_{T\beta_0}(j)}{\partial\beta} &= -\left(\frac{1}{T}\sum_{t=1+j}^{T}\left(\frac{Y_{t-\tau}-\lambda_{t-j}}{\lambda_t\lambda_{t-j}}\right)\mathbf{X}'_t + \frac{1}{T}\sum_{t=1+j}^{T}\left(\frac{Y_t-\lambda_t}{\lambda_t\lambda_{t-j}}\right)\mathbf{X}'_{t-j}\right)\\ &= -\mathbf{A}_{T,1} - \mathbf{A}_{T,2}\end{aligned}$$

where

$$\mathbf{A}_{T,1} = \frac{1}{T}\sum_{t=1+j}^{T}\left(\frac{Y_{t-j}-\lambda_{t-j}}{\lambda_t\lambda_{t-j}}\right)\mathbf{X}'_t;\quad \mathbf{A}_{T,2} = \frac{1}{T}\sum_{t=1+\tau}^{T}\left(\frac{Y_t-\lambda_t}{\lambda_t\lambda_{t-j}}\right)\mathbf{X}'_{t-j}.$$

Setting $\boldsymbol{\zeta}^{(i)}_{\beta_0}(j) := \lim_{T\to\infty} E[\mathbf{A}_{T.i}(j)], i = 1,2$, we wish to show that $\mathbf{A}_{T,i}(j) = \boldsymbol{\zeta}^{(i)}_{\beta_0}(j) + o_p(1), i = 1,2;\ j = 1,2,3,\ldots$.

It suffices to show that $E\left\|\mathbf{A}_{T,i}(j) - E\left[\mathbf{A}_{T,i}(j)\right]\right\|^2$ is $o(1)$, $i = 1,2$. First consider $i = 1$,

$$E\left\|A_{T,i}(j) - E\left[A_{T,i}(j)\right]\right\|^2 = \frac{1}{T^2}\sum_{t=j+1}^{T}\sum_{r=j+1}^{T} E\left[e(t,t-j)'e(r,r-j)\right] = o(1) \tag{14}$$

where $e(t,t-j) = (Y_{t-j}-\lambda_{t-j})\left(\lambda_t\lambda_{t-j}\right)^{-1}\mathbf{X}'_t - E\left[(Y_{t-j}-\lambda_{t-j})\left(\lambda_t\lambda_{t-j}\right)^{-1}\mathbf{X}'_t\right]$ and, henceforth we omit dependence on β_0 in the notation.

For some $n > 0$ fixed with $T, E\left\|\mathbf{A}_{T,1}(j) - E\left[\mathbf{A}_{T,1}(j)\right]\right\|^2$ is

$$\begin{aligned}\frac{1}{T^2}\sum_{t=j+1}^{T} E\left[e(t,t-j)'e(t,t-j)\right] + \frac{2}{T^2}\sum_{t=j+1}^{T}\sum_{t-n-j\le r<t}^{T} E\left[e(t,t-j)'e(r,r-j)\right] \qquad (15)\\ + \frac{2}{T^2}\sum_{t=j+1}^{T}\sum_{j+1\le r<t-n-j}^{T} E\left[e(t,t-j)'e(r,r-j)\right].\end{aligned}$$

The first two terms of (15) are $O(T^{-1}) = o(1)$ since it involves a maximum of $T+n$ terms with bounded absolute expectation, since by Assumptions 1-2 and Minkowski and

Cauchy-Schwarz inequalities,

$$
\begin{aligned}
& E\left\| \left(\frac{Y_{t-j}-\lambda_{t-j}}{\lambda_t\lambda_{t-j}}\right)\mathbf{X}_t' - E\left[\left(\frac{Y_{t-j}-\lambda_{t-j}}{\lambda_t\lambda_{t-j}}\right)\mathbf{X}_t'\right]\right\|^2 \\
& \leq 2E\left\|\left(\frac{Y_{t-j}-\lambda_{t-j}}{\lambda_t\lambda_{t-j}}\right)\mathbf{X}_t'\right\|^2 \\
& = 2E\left\|\frac{\varepsilon_{t-j}-1}{\lambda_t}\mathbf{X}_t'\right\|^2 \\
& \leq 2\left(E\left\|\frac{1}{\lambda_t}\mathbf{X}_t'\right\|^4\right)^{1/2}\left(E|\varepsilon_t-1|^4\right)^{1/2} < \infty
\end{aligned}
$$

In order to show that the third term of (15) is bounded, notice that $e(r, r-j)$ is $\mathcal{F}_1^r$ measurable and that $e(t, t-j)$ is $\mathcal{F}_t^\infty$ measurable. Given Assumption 2, $E\left\|e(t,t-j)\right\|^{2+\delta} < \infty$, $E\left\|e(r,r-j)\right\|^{2+\delta} < \infty$, we can use Roussas and Ioannides (1987) moment inequality to show that the third term of (15) is bounded in absolute value by

$$
\frac{C}{T^2}\left(E\left\|e(t,t-j)\right\|^{2+\delta}E\left\|e(r,r-j)\right\|^{2+\delta}\right)^{2+\delta}\sum_{t=j+1}^{T}\sum_{j+1\leq r<t-n-j}^{T}\alpha_{t-j-r}^{\frac{\delta}{2+\delta}} = O(T^{-1}) = o(1).
$$

Using exactly the same procedure, we can show that $E\left\|\mathbf{A}_{T,2}(j) - E\left[\mathbf{A}_{T,2}(j)\right]\right\|^2$ is $o(1)$. Then, we have that under H_0

$$
\frac{\partial\gamma_{T\beta_0}(j)}{\partial\beta} = -E\left(\frac{\varepsilon_{t-j}-1}{\lambda_t}\mathbf{X}_t'\right) - E\left(\frac{\varepsilon_t-1}{\lambda_{t-j}}\mathbf{X}_{t-j}'\right) + o_p(1).
$$

Now, we just need to show that the second order term on the expansion is $o_p(T^{-1/2})$. In order to do that, we just need to show that $\ddot{\boldsymbol{\rho}}_{T\beta_{T,j}^*}(j) = \partial^2\hat{\rho}_{T\beta}(j)/\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}'$ is $O_p(1)$. For $j = 1, \ldots, m$ we have

$$
\begin{aligned}
\ddot{\rho}_{T\beta_{T,j}^*}(j) = {} & \frac{\frac{\partial^2}{\partial\beta\partial\beta'}\hat{\gamma}_{T\beta^*}(j)}{\hat{\gamma}_{T\beta^*}(0)} - \frac{\frac{\partial}{\partial\beta}\hat{\gamma}_{T\beta^*}(j)}{\hat{\gamma}_{T\beta^*}(0)}\frac{\left(\frac{\partial}{\partial\beta}\hat{\gamma}_{T\beta^*}(j)\right)'}{\hat{\gamma}_{T\beta^*}(0)} - \frac{\partial}{\partial\beta'}\hat{\rho}_{T\beta^*}(j)\frac{\frac{\partial}{\partial\beta}\hat{\gamma}_{T\beta^*}(0)}{\hat{\gamma}_{T\beta^*}(0)} \\
& - \hat{\rho}_{T\beta^*}(j)\left(\frac{\frac{\partial^2}{\partial\beta\partial\beta'}\hat{\gamma}_{T\beta^*}(0)}{\hat{\gamma}_{T\beta^*}(0)} - \frac{\frac{\partial}{\partial\beta}\hat{\gamma}_{T\beta^*}(0)}{\hat{\gamma}_{T\beta^*}(0)}\frac{\left(\frac{\partial}{\partial\beta}\hat{\gamma}_{T\beta^*}(0)\right)'}{\hat{\gamma}_{T\beta^*}(0)}\right),
\end{aligned}
$$

where

$$\frac{\partial^2}{\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}'}\hat{\gamma}_{T\beta^*}(j) = -\frac{\partial}{\partial\boldsymbol{\beta}'}\boldsymbol{A}_{T,1}\bigg|_{\boldsymbol{\beta}=\boldsymbol{\beta}^*} - \frac{\partial}{\partial\boldsymbol{\beta}'}\boldsymbol{A}_{T,2}\bigg|_{\boldsymbol{\beta}=\boldsymbol{\beta}^*}$$

$$\frac{\partial^2}{\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}'}\hat{\gamma}_{T\beta^*}(0) = -2\frac{\partial}{\partial\boldsymbol{\beta}'}\boldsymbol{B}_T\bigg|_{\boldsymbol{\beta}=\boldsymbol{\beta}^*}$$

and

$$\boldsymbol{B}_T = \frac{1}{T}\sum_{t=1}^{T}\left(\frac{Y_t-\lambda_t}{\lambda_t^2}\right)\mathbf{X}_t'$$

$$\frac{\partial}{\partial\boldsymbol{\beta}'}\boldsymbol{A}_{T,1}\bigg|_{\boldsymbol{\beta}=\boldsymbol{\beta}^*} = -\frac{1}{T}\sum_{t=j+1}^{T}\left(\mathbf{X}_t\frac{Y_{t-j}-\lambda_{t-j}^*}{\lambda_t^*\lambda_{t-j}^*}\mathbf{X}_t' + \mathbf{X}_{t-j}\frac{Y_{t-j}}{\lambda_t^*\lambda_{t-j}^*}\mathbf{X}_t'\right)$$

$$\frac{\partial}{\partial\boldsymbol{\beta}'}\boldsymbol{A}_{T,2}\bigg|_{\boldsymbol{\beta}=\boldsymbol{\beta}^*} = -\frac{1}{T}\sum_{t=j+1}^{T}\left(\mathbf{X}_{t-j}\frac{Y_t-\lambda_t^*}{\lambda_t^*\lambda_{t-j}^*}\mathbf{X}_{t-j}' + \mathbf{X}_t\frac{Y_t}{\lambda_t^*\lambda_{t-j}^*}\mathbf{X}_{t-j}'\right)$$

$$\frac{\partial}{\partial\boldsymbol{\beta}'}\boldsymbol{B}_T\bigg|_{\boldsymbol{\beta}=\boldsymbol{\beta}^*} = \frac{-2}{T}\sum_{t=1}^{T}\mathbf{X}_t\frac{Y_t-\lambda_t^*}{\lambda_t^{2*}}\mathbf{X}_t'$$

and $\lambda_s^* = \exp(\boldsymbol{X'}_s\boldsymbol{\beta}^*)$.

Using Assumptions (1) and (2) and techniques similar to the ones we already used, we can show that $\ddot{\rho}_{T\beta^*_{T,j}}(j) = O_p(1)$. ■

PROOF OF THEOREM 1

Using algebra and Proposition 1, we find that $\hat{\mathcal{L}}_T^{(m)}\tilde{\boldsymbol{\rho}}_{T\hat{\beta}_T}^{(m)} = \hat{\mathcal{L}}_T^{(m)}\tilde{\boldsymbol{\rho}}_{T\beta_0}^{(m)} + o_p(T^{-1/2})$, because from Assumption 3,

$$\tilde{\boldsymbol{\theta}}_{T\hat{\beta}_T}^{(\tau)}[\bar{\boldsymbol{\rho}}_{T\beta_T}^{(m)}] = \tilde{\boldsymbol{\theta}}_{T\hat{\beta}_T}^{(\tau)}[\bar{\boldsymbol{\rho}}_{T\beta_0}^{(m)}] + \left(\hat{\boldsymbol{\beta}}_T - \boldsymbol{\beta}\right) + o_p(T^{-1/2}),$$

$\tau = 1,\ldots,m-k$, such that $\tilde{\boldsymbol{\theta}}_{T\beta}^{(\tau)}[\boldsymbol{\rho}_{T\beta}] = \left(\sum_{l=\tau+1}^{m}\hat{\boldsymbol{\xi}}_{T\beta}(l)'\hat{\boldsymbol{\xi}}_{T\beta}(l)\right)^{-1}\sum_{l=\tau+1}^{m}\hat{\boldsymbol{\xi}}_{T\beta}(l)'\rho_{T\beta}(l)$ and $\hat{\boldsymbol{\xi}}_{T\beta_T}(\tau) \to_p \boldsymbol{\xi}_{\beta_0}(\tau)$, which can be proved using the same methods used in the proof of Proposition 1.

Similar, we can show that$\hat{\mathcal{L}}_T^{(m)}\tilde{\rho}_{T\beta_0}^{(m)}(\tau) = \tilde{\rho}_{T\beta_0}^{(m)}(\tau) - \boldsymbol{\xi}_{\beta_0}(\tau)\boldsymbol{\theta}_{\beta_0}^{(\tau)}[\tilde{\rho}_{T\beta_0}^{(m)}(\tau)] + o_p(T^{-1/2})$,where $\boldsymbol{\theta}_{\beta}^{(\tau)}[\rho] = \left(\sum_{l=\tau+1}^{m}\boldsymbol{\xi}_\beta(l)'\boldsymbol{\xi}_\beta(l)\right)^{-1}\sum_{l=\tau+1}^{m}\boldsymbol{\xi}_\beta(l)'\rho(l), \tau = 1,\ldots,m-k.$

The CLT for $\bar{\boldsymbol{\rho}}_{T\beta_0}^{(m)}$ follows from the CLT for $\tilde{\boldsymbol{\rho}}_{T\beta_0}^{(m)}$ under Assumptions 1, condition (3), H_0 and from the fact that $\bar{\boldsymbol{\rho}}_{T\beta_0}^{(m)}$ are standardized by construction if $\tilde{\boldsymbol{\rho}}_{T\beta_0}^{(m)}$ is already standardized.

Under H_0, $\tilde{\rho}^{(m)}_{T\beta_0}(\tau) = 0$ for all $\tau = 1, \ldots, m$ and hence $\bar{\boldsymbol{\rho}}^{(m)}_{T\beta_0}$ has asymptotic mean equal to 0. In order to show the asymptotic variance of $\bar{\boldsymbol{\rho}}^{(m)}_{T\beta_0}$ is equal to I_{m-k}, notice that $AVar\left(T^{1/2}\hat{\mathcal{L}}^{(m)}_T\tilde{\rho}^{(m)}_{T\beta_0}(\tau)\right)$ is equal to

$$AVAR\left(T^{1/2}\left(\tilde{\rho}^{(m)}_{T\beta_0}(\tau) - \hat{\boldsymbol{\xi}}_{T\beta_T}(\tau)\tilde{\theta}^{(\tau)}_{T\hat{\beta}_T}\right)\right) = AVAR\left(T^{1/2}\left(\tilde{\rho}^{(m)}_{T\beta_0}(\tau) - \boldsymbol{\xi}_{\beta_0}(\tau)\boldsymbol{\theta}^{(\tau)}_{\beta_0}\left[\tilde{\boldsymbol{\rho}}^{(m)}_{T\beta_0}\right]\right)\right)$$
$$= 1 + \boldsymbol{\xi}_{\beta_0}(\tau)\left(\sum_{l=\tau+1}^{m}\boldsymbol{\xi}_{\beta_0}(l)'\boldsymbol{\xi}_{\beta_0}(l)\right)^{-1}\boldsymbol{\xi}_{\beta_0}(\tau)',$$

while for $1 \leq \tau < q \leq m-k$, $ACov\left(T^{1/2}\hat{\mathcal{L}}^{(m)}_T\tilde{\rho}^{(m)}_{T\beta_0}(\tau), T^{1/2}\hat{\mathcal{L}}^{(m)}_T\tilde{\rho}^{(m)}_{T\beta_0}(q)\right)$ is given by

$$\begin{aligned}
&ACov\left(T^{1/2}\left(\tilde{\rho}^{(m)}_{T\beta_0}(\tau) - \hat{\boldsymbol{\xi}}_{T\beta_T}(\tau)\tilde{\boldsymbol{\theta}}^{(\tau)}_{T\hat{\beta}_T}\right), T^{1/2}\left(\tilde{\rho}^{(m)}_{T\beta_0}(q) - \hat{\boldsymbol{\xi}}_{T\beta_T}(q)\tilde{\boldsymbol{\theta}}^{(q)}_{T\hat{\beta}_T}\right)\right)\\
&= ACov\left(T^{1/2}\left(\tilde{\rho}^{(m)}_{T\beta_0}(\tau) - \boldsymbol{\xi}_{\beta_0}(\tau)\boldsymbol{\theta}^{(\tau)}_{\beta_0}\left[\tilde{\boldsymbol{\rho}}^{(m)}_{T\beta_0}\right]\right), T^{1/2}\left(\tilde{\rho}^{(m)}_{T\beta_0}(q) - \boldsymbol{\xi}_{\beta_0}(q)\boldsymbol{\theta}^{(q)}_{\beta_0}\left[\tilde{\boldsymbol{\rho}}^{(m)}_{T\beta_0}\right]\right)\right)\\
&= ACov\left(T^{1/2}\tilde{\rho}^{(m)}_{T\beta_0}(\tau), T^{1/2}\tilde{\rho}^{(m)}_{T\beta_0}(q)\right) - ACov\left(T^{1/2}\tilde{\rho}^{(m)}_{T\beta_0}(\tau), T^{1/2}\boldsymbol{\xi}_{\beta_0}(q)\boldsymbol{\theta}^{(q)}_{\beta_0}\left[\tilde{\boldsymbol{\rho}}^{(m)}_{T\beta_0}\right]\right)\\
&- ACov\left(T^{1/2}\boldsymbol{\xi}_{\beta_0}(\tau)\boldsymbol{\theta}^{(\tau)}_{\beta_0}\left[\tilde{\boldsymbol{\rho}}^{(m)}_{T\beta_0}\right], T^{1/2}\tilde{\rho}^{(m)}_{T\beta_0}(q)\right)\\
&+ ACov\left(T^{1/2}\boldsymbol{\xi}_{\beta_0}(\tau)\boldsymbol{\theta}^{(\tau)}_{\beta_0}\left[\tilde{\boldsymbol{\rho}}^{(m)}_{T\beta_0}\right], T^{1/2}\boldsymbol{\xi}_{\beta_0}(q)\boldsymbol{\theta}^{(q)}_{\beta_0}\left[\tilde{\boldsymbol{\rho}}^{(m)}_{T\beta_0}\right]\right)
\end{aligned}$$

where the terms are respectively equal to $0, 0$, $-\boldsymbol{\xi}_{\beta_0}(\tau)\left(\sum_{l=\tau+1}^{m}\boldsymbol{\xi}_{\beta_0}(l)'\boldsymbol{\xi}_{\beta_0}(l)\right)^{-1}\boldsymbol{\xi}_{\beta_0}(q)'$ and $\boldsymbol{\xi}_{\beta_0}(\tau)\left(\sum_{l=\tau+1}^{m}\boldsymbol{\xi}_{\beta_0}(l)'\boldsymbol{\xi}_{\beta_0}(l)\right)^{-1}\boldsymbol{\xi}_{\beta_0}(q)'$, and the asymptotic covariance of the projection is 0. ■

PROOF OF THEOREM 2

The result follows from noticing that under H_{1T} Proposition 1 is still valid, because for each $j = 1, \ldots, m$, we have

$$\begin{aligned}
\frac{\partial}{\partial\beta}\hat{\rho}_{T\beta}(j) &= \frac{\frac{\partial}{\partial\beta}\hat{\gamma}_{T\beta}(j)}{\hat{\gamma}_{T\beta}(0)} - \frac{\hat{\gamma}_{T\beta}(j)}{\hat{\gamma}_{T\beta}(0)}\frac{\frac{\partial}{\partial\beta}\hat{\gamma}_{T\beta}(0)}{\hat{\gamma}_{T\beta}(0)}\\
&= \frac{\frac{\partial}{\partial\beta}\hat{\gamma}_{T\beta}(j)}{\hat{\gamma}_{T\beta}(0)} - \frac{r(j)}{\sqrt{T}}O_p(1) + O_p(T^{-1})\\
&= \frac{\frac{\partial}{\partial\beta}\hat{\gamma}_{T\beta}(j)}{\hat{\gamma}_{T\beta}(0)} + O_p(T^{-1/2}),
\end{aligned}$$

and then, for each $j = 1, \ldots, m$, we have

$$\begin{aligned}\hat{\rho}^{(m)}_{T\hat{\beta}_T}(j) - \hat{\rho}^{(m)}_{T\beta_0}(j) &= \frac{\frac{\partial}{\partial\beta}\hat{\gamma}_{T\beta}(j)}{\hat{\gamma}_{T\beta}(0)}(\hat{\beta}_T - \beta_0) + O_p(T^{-1/2})(\hat{\beta}_T - \beta_0) + o_p(T^{-1/2}) \\ &= \frac{\frac{\partial}{\partial\beta}\hat{\gamma}_{T\beta}(j)}{\hat{\gamma}_{T\beta}(0)}(\hat{\beta}_T - \beta_0) + o_p(T^{-1/2}).\end{aligned}$$

Hence, from Theorem 1, we have that

$$\begin{aligned}\hat{\mathcal{L}}^{(m)}_T \tilde{\rho}^{(m)}_{T\hat{\beta}_T}(\tau) &= \hat{\mathcal{L}}^{(m)}_T \tilde{\rho}^{(m)}_{T\beta_0}(\tau) + o_p(T^{-1/2}) \\ &= \tilde{\rho}^{(m)}_{T\beta_0}(\tau) - \boldsymbol{\xi}_{\beta_0}(\tau)\boldsymbol{\theta}^{(\tau)}_{\beta_0}[\tilde{\rho}^{(m)}_{T\beta_0}(\tau)] + o_p(T^{-1/2}),\end{aligned}$$

$\tau = 1, \ldots, m-k$, also under H_{1T}.

We have seem in Theorem 1 that, under Assumptions 1 - 3, the CLT for $\bar{\boldsymbol{\rho}}^{(m)}_{T\beta_0}$ follows from the CLT for $\tilde{\boldsymbol{\rho}}^{(m)}_{T\beta_0}$. Since under H_{1T} $\tilde{\boldsymbol{\rho}}^{(m)}_{T\beta_0}$ has asymptotic mean equal to $\tilde{\boldsymbol{h}}^{(m)}_{\beta_0} = \left(h^{(m)}_{\beta_0}(1), \ldots, h^{(m)}_{\beta_0}(m)\right)'$, with $h^{(m)}_{\beta}(\tau)$ as in (13), it is clear that, for $\tau = 1, \ldots, m-k$, $\hat{\mathcal{L}}^{(m)}_T \tilde{\rho}^{(m)}_{T\beta_0}(\tau)$ is asymptotically normal, with an asymptotic drift equal to $\check{h}^{(m)}_{\beta_0}(\tau)$, defined in (12), and asymptotic variance equal to $1 + \boldsymbol{\xi}_{\beta_0}(\tau)\left(\sum_{l=\tau+1}^{m} \boldsymbol{\xi}_{\beta_0}(l)'\boldsymbol{\xi}_{\beta_0}(l)\right)^{-1}\boldsymbol{\xi}_{\beta_0}(\tau)'$.

Since $\hat{\mathcal{L}}^{(m)}_T \tilde{\rho}^{(m)}_{T\beta_0}(\tau)$ is asymptotic independent of $\hat{\mathcal{L}}^{(m)}_T \tilde{\rho}^{(m)}_{T\beta_0}(q)$,for $1 \leq \tau < q \leq m-k$, as shown in Theorem 1, the result follows. ■

References

Benjamin, M. A., Rigby, R. A., and Stasinopoulos, D. M. (2003), "Generalized autoregressive moving average models," *Journal of the American Statistical Association*, 98, 214–223.

Box, G., and Pierce, D. A. (1970), "Distribution of residual autocorrelations in autoregressive-integrated moving average time series models," *Journal of the American Statistical Association*, 65, 1509–1526.

Brown, R. L., Durbin, J., and Evans, J. M. (1975), "Techniques for testing the constancy of regression relationships over time," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 37, 149–192.

Das, S. R., Duffie, D., Kapadia, N., and Saita, L. (2007), "Common failings: how corporate defaults are correlated," *The Journal of Finance*, 62, 93–118.

Davis, R. A., Dunsmuir, W. T., and Streett, S. B. (2003), "Observation-driven models for Poisson counts," *Biometrika*, 90, 777–790.

Davis, R. A., Dunsmuir, W. T., and Wang, Y. (2000), "On autocorrelation in a Poisson regression model," *Biometrika*, 87, 491–505.

Delgado, M. A., and Velasco, C. (2010), "Distribution-free tests for time series models specification," *Journal of Econometrics*, 155, 128–137.

Delgado, M. A., and Velasco, C. (2011), "An asymptotically pivotal transform of the residuals sample autocorrelations with application to model checking," *Journal of the American Statistical Association*, 116, 946–958.

Duffie, D., Eckner, A., Horel, G., and Saita, L. (2009), "Frailty correlated default," *The Journal of Finance*, 64, 2089–2123.

Escanciano, J. C. (2009), "On the lack of power of omnibus specification tests," *Econometric Theory*, 25, 162.

Escanciano, J. C., and Lobato, I. N. (2009), "An automatic Portmanteau test for serial correlation," *Journal of Econometrics*, 151, 140–149.

Escanciano, J. C., Lobato, I. N., and Zhu, L. (2013), "Automatic specification testing for vector autoregressions and multivariate nonlinear time series models," *Journal of Business & Economic Statistics*, (Forthcoming).

Fokianos, K. (2012), "Count time series," *in: Subba Rao, T. , S. Subba Rao and C.R. Rao, (Eds.), Handbook of Statistics, Vol.30. Elsevier, Amsterdam, pp. 315–348*, .

Fokianos, K., and Neumann, M. H. (2013), "A goodness-of-fit test for Poisson count processes," *Eletronic Journal of Statistics*, 7, 793–819.

Fokianos, K., and Tjøstheim, D. (2011), "Log-linear Poisson autoregression," *Journal of Multivariate Analysis*, 102, 563–578.

Francq, C., Roy, R., and Zakoïan, J. M. (2005), "Diagnostic checking in ARMA models with uncorrelated errors," *The Journal of the American Statistical Association*, 100, 532–544.

Hong, Y. (1996), "Consistent testing for serial correlation of unknown form," *Econometrica*, 64, 837.

Jung, R. C., and Tremayne, A. R. (2003), "Testing for serial dependence in time series models of counts," *Journal of Time Series Analysis*, 24, 65–84.

Koopman, S. J., Lucas, A., and Schwaab, B. (2011), "Modeling frailty-correlated defaults using many macroeconomic covariates," *Journal of Econometrics*, 162, 312–325.

Koopman, S. J., Lucas, A., and Schwaab, B. (2012), "Dynamic factor models with macro, frailty, and industry effects for U.S. default counts: the credit crisis of 2008," *Journal of Business & Economic Statistics*, 30, 521–532.

Ljung, G. M., and Box, G. E. P. (1978), "On a measure of lack of fit in time series models," *Biometrika*, 65, 297.

Lobato, I. N., Nankervis, J. C., and Savin, N. E. (2002), "Testing for zero autocorrelation in the presence of statistical dependence," *Econometric Theory*, 18, 730–743.

Neumann, M. H. (2011), "Absolute regularity and ergodicity of Poisson count processes," *Bernoulli*, 17, 1268 – 1284.

Romano, J., and Thombs, L. (1996), "Inference for autocorrelations under weak assumptions," *Journal of the American Statistical Association*, 91, 590–600.

Roussas, G. G., and Ioannides, D. (1987), "Moment inequalities for mixing sequences of random variables," *Stochastic Analysis and Applications*, 5, 61–120.

Sun, J., and McCabe, B. (2013), "Score statistics for testing the serial dependence in count data score statistics for testing the serial dependence in count data," *Journal of Time Series Analysis*, 34, 315–329.

Tarullo, D. K. (2010), “Equipping financial regulators with the tools necessary to monitor systemic risk,” *Statement before the banking subcommittee on security and international trade and finance U.S. senate, 12 February 2010*, .

Tjøstheim, D. (2012), “Some recent theory for autoregressive count time series,” *TEST*, 21, 413–438.

Zeger, S. L. (1988), “A regression model for time Series of counts,” *Biometrika*, 75, 621–629.

Zeger, S. L., and Qaqish, B. (1988), “Markov regression models for time series: a quasi-likelihood approach.,” *Biometrics*, 44, 1019–31.

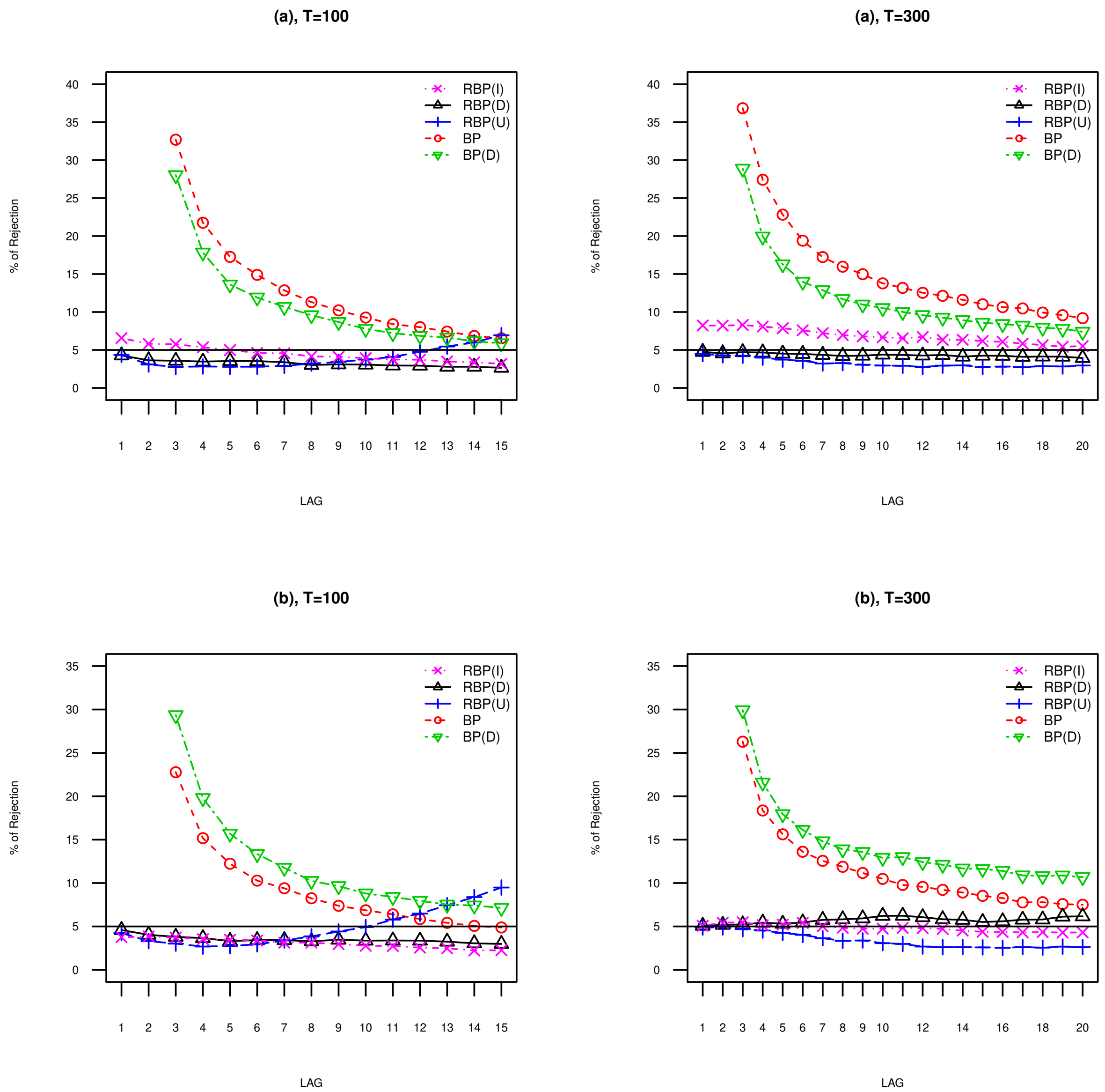

Figure 1: Empirical size of Portmanteau tests at 5 % significance. RBP(W) is $\bar{B}^{(m)}_{T\hat{\beta}_T}(s)$ based on recursive projected residuals autocorrelations, using $\hat{A}^m_t = W$, compared with a $\chi^2_{(s)}$ critical value. $W = I$ means $\hat{A}^m_t = I_m$, $W = D$ means $\hat{A}^m_t$ is diagonal, and $W = U$ means an unrestricted estimate of $\hat{A}^m_t$ is used. BP and BP(D) are the classical Box-Pierce test and the standardized Box-Pierce test, respectively, compared with a $\chi^2_{(s-2)}$ critica value.

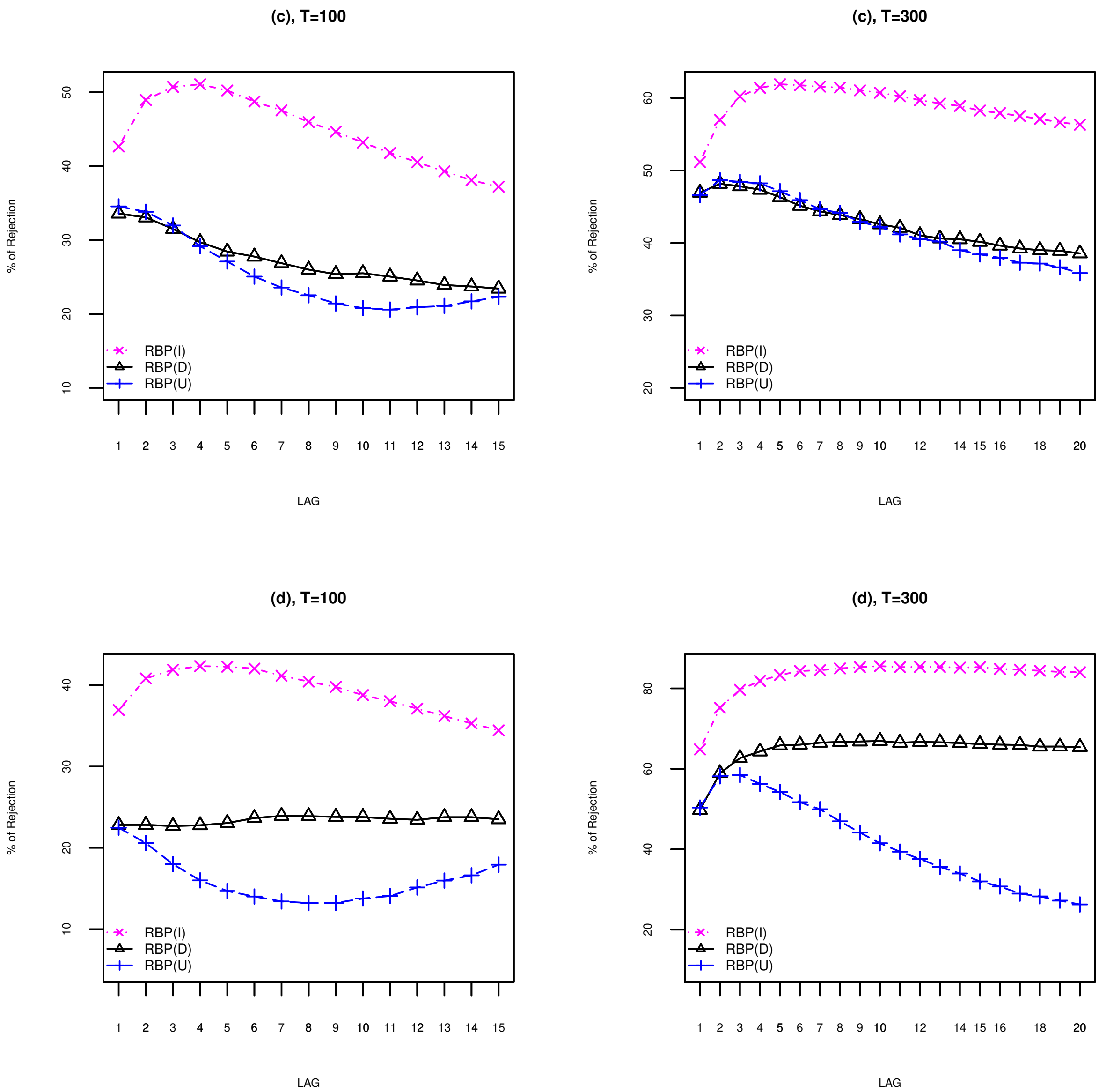

Figure 2: Empirical Power of Portmanteau tests at 5 % significance. RBP(W) is $\bar{B}^{(m)}_{T\hat{\beta}_T}(s)$ based on recursive projected residuals autocorrelations, using $\hat{A}^m_t = W$, compared with a $\chi^2_{(s)}$ critical value. $W = I$ means $\hat{A}^m_t = I_m$, $W = D$ means $\hat{A}^m_t$ is diagonal, and $W = U$ means an unrestricted estimate of $\hat{A}^m_t$ is used.

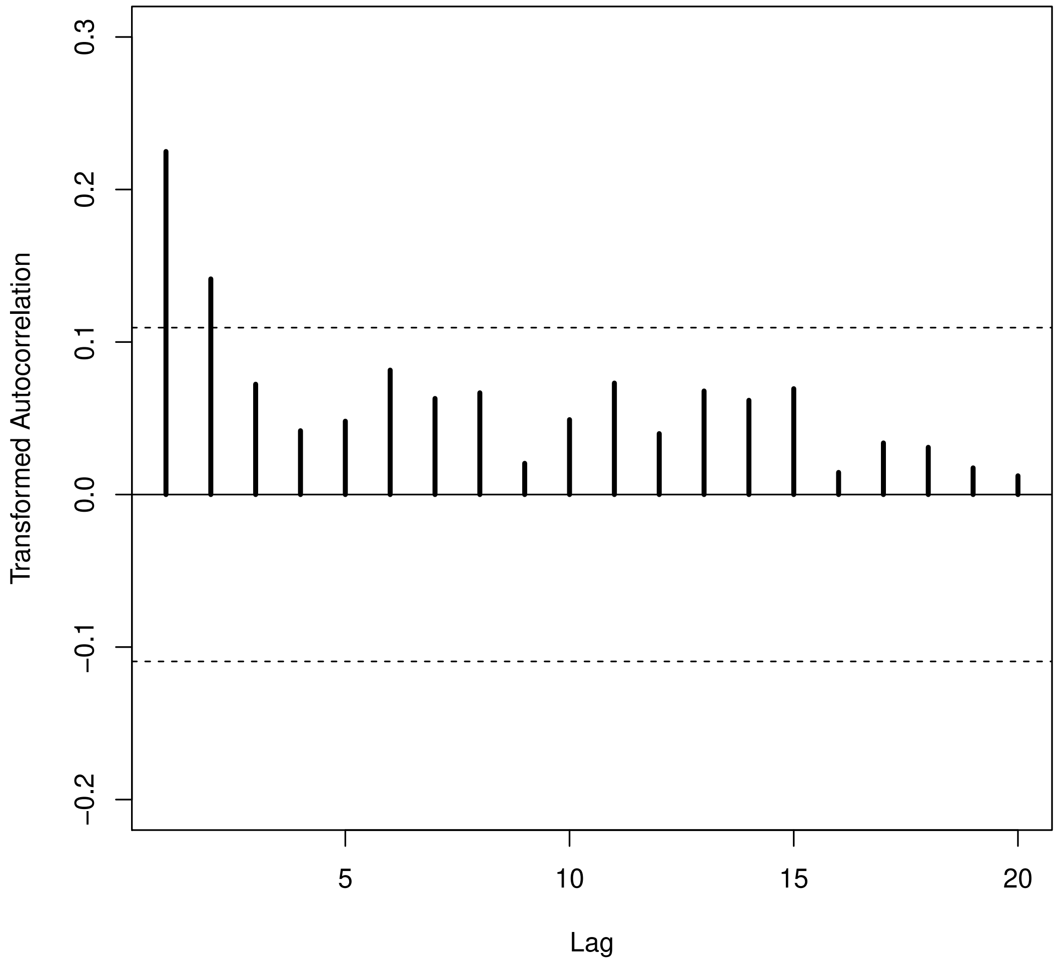

Figure 3: Projected Sample Autocorrelations from Poisson Regression Residuals

Table 1: Macroeconomic Time Series data

Variable	Shortname
Industrial production index	INDPRO
New housing permits	PERMIT
Civilian unemployment rate	UNRTAE
Moody's Seasoned Baa Corporate Bond Yield	BAA
10-Year Treasury Constant Maturity Rate	GS10
Federal Funds Rate	FEDFUNDS
Producer Price Index: Finished Goods	PPIFGS
Producer Price Index: Fuels and related energy	PPIENG
S&P 500 yearly returns	SP500RET
S&P 500 return volatility	SP500VOL
2005 Bankruptcy Act	DUMMY2005

Table 2: Goodness-of-Fit Analysis for U.S.A. Bankruptcy Counts based on Poisson model

Models	$\hat{A}_t^m$ / s	1	2	3	4	5	6
Poisson	I_m	15.641***	19.65***	25.193***	24.228***	25.904***	26.379***
Static	$diag$	11.123***	14.109***	17.328***	16.989***	18.948***	19.261***
Poisson	I_m	0.069	1.657	2.085	2.261	2.616	4.525
GARMA(0,1)	$diag$	0.065	1.863	2.462	2.444	2.778	4.945
Poisson	I_m	0.001	0.192	0.242	0.223	0.237	1.587
GARMA(1,0)	$diag$	0.001	0.202	0.254	0.225	0.234	1.605

Note: ***denotes significant at 1% level. The test statistic is $\bar{B}^{(m)}_{T\hat{\beta}_T}(s)$, with $\hat{A}_t^m$ equal to either the Identity or Diagonal matrix. $m = 12 + s$ for the Static model and $m = 13 + s$ for the GARMA models

Table 3: Estimated parameters of Poisson models for U.S.A. Bankruptcy Counts

COVARIATES	STATIC	GARMA(1,0)	GARMA(0,1)
INTERCEPT	1.263***	1.281***	1.29***
Dummy2005	-0.303*	-0.242	-0.27**
BAA	0.270***	0.173*	0.189*
FEDFUNDS	-0.015	-0.034	-0.034
GS10	-0.11	-0.047	-0.054
INDPRO	-0.063***	-0.084***	-0.083***
PERMIT	0.017***	0.012***	0.014***
PPIENG	0.028***	0.026***	0.028***
PPIFGS	-0.112***	-0.113***	-0.118***
SP500RET	-0.003	-0.004	-0.004
SP500VOL	-0.001	0.000	-0.001
UNRATE	0.341***	0.171	0.199***
AR.1		0.334***	
MA.1			0.307***
BIC	1361.759	1306.185	1312.192
AIC	1316.025	1256.679	1262.686

Note: *,**,*** denote significant at 10%, 5% and 1% level.