Count Data in Time Series

Koen Bel
311818

Econometrics and Operational Research
Major Econometrics

Bachelor Thesis 2010

July 2, 2010

Abstract

This research paper is about time series of count data. The common methods, like the Autoregressive Moving Average model, are not profitable to use, because they do not account for integer-valued positive numbers. That is why other methods will be proposed. The Integer-valued Autoregressive model and the newly suggested Autoregressive Conditional Integer-valued model will be explained and compared. This article concludes three things: the theoretically best method to use with count data is the ACI-model, because of few violated assumptions, such as possibly negative correlations and the assumption of the presence of a Poisson distribution; the best method on the basis of a simulation study is the INAR-model, shown by the fact that this model is even doing good when data is simulated by another method; the best method for particular purchase data is the Poisson-model, but that is mostly due to the low existence of autocorrelation. All in all, the INAR-model seems most useful for time series of count data.

Contents

1 Introduction

For time series certain models are commonly known, like an Autoregressive Moving Average, or ARMA, model. This method takes into account that past information influences the variables of today. For example, the rate of inflation of a few days ago influences the rate of inflation today. Or import and export data can be predicted by using information on how many is imported and exported in the past few months. Many examples can by given and these ARMA-models work pretty well in modeling the time series.

A problem occurs when count data have to be modeled. The number of counts in a certain period can only be an integer and that is why the commonly used ARMA-model, which assumes real numbers, seems not very useful anymore. For this integer-valued regressions, methods can be used which make use of the Poisson Distribution. These methods will be proposed in this paper. The methods can be used for more than a few practical situations. For example, the number of traffic accidents in the past can be used to predict or prevent accidents in the future (Quddus, 2008, p. 1734). An other example is about the number of purchases of a product in the near past which can be used to model the purchases of a store (Böckenholt, 1999, p. 328).

The goal of this paper is to find out which method can be used best for modeling count data in time series. Different methods will be proposed and compared, like the ARMA-model, the Integer-valued Autoregressive, or INAR, model and the Autoregressive Conditional Integer-valued, or ACI, model. In Section 2 the methods appropriate to use are proposed and explained. Furthermore, this section looks at the assumptions made by the different methods. In Section 3 it will be shown how a method can be selected. This will be done by simulating data such that a comparison can be made of which kind of data can be modeled by which method. In Section 4 purchase data of five households are modeled, by using the appropriate methods. At the end of this paper, a conclusion is drawn about which method suits best for time series of count data.

2 Model Proposals

For time series of count data different types of methods can be used. First of all, this paper will present the ARMA-model, which is commonly used for time series, but probably not good enough to handle integer-valued variables. The ARMA-model has the assumption that the error term is normally distributed and with count data, this assumption seems not very appropriate. That is why other methods will be used too: the Poisson-model, the INAR-model and the ACI-model. The latter is a method which is not yet proposed in other papers. In this section these methods will be explained and it will be shown why these methods are applicable for time series of count data.

An important question to be answered is which assumptions are in line with time series of count data and which assumptions are not. When a method has many assumptions which are not in line with count data, this method would theoretically be not very useful for this kind of data. This section also takes a look at these assumptions.

2.1 ARMA(p, q)-model

In the Autoregressive Moving Average model two different components can be used to model a time series: an autoregressive and a moving average part. The first takes into account relations to previous observations: y_t is correlated with $y_{t-1}, ..., y_{t-p}$, where y_t is the counted observation at time t, while the second part correlates previous error terms with the current observation. This model is defined as

$$y_t = \alpha_t + \psi_1 y_{t-1} + ... + \psi_p y_{t-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + ... + \theta_q \epsilon_{t-q}, \tag{1}$$

where α_t, ψ_i and θ_i are the parameters and ϵ_t is the error term, which is $N(0, \sigma)$ distributed (Franses and Van Dijk, 2009, p. 60-62). By choosing the numbers for p and q, one chooses how many lags should be used in the ARMA(p, q)-model. From now on, p and q will be equal to 1, for simplicity.

This ARMA$(1, 1)$-model can be extended by adding a trend (upward or downward) or adding seasonal effects. Also, aberrant observations can be detected and non-linearity can be solved (see for a detailed description Franses and Van Dijk, 2009). This method can even account for the restriction of positive numbers by changing y_t in $\ln(1 + y_t)$. Furthermore, this method has some assumptions. The mean of the error terms should be equal to zero, all error terms should have the same variance and their is no correlation between the error terms. When these three assumptions hold, the error term in the model is said to be 'white noise' (Franses and Van Dijk, 2009, p. 51-52). For count data, it is not harmful to have these assumptions. However, the residuals in a ARMA-model should be approximately normally distributed. This assumptions is violated when using count data, because negative observations cannot occur. The symmetric form then will not be found, because negative values can be predicted, while only positive values should be found. Lastly, the ARMA-model does not take into account the fact that data is only integer-valued and this missing fact makes this method less appropriate.

2.2 Poisson model

Because integer-valued data should be accounted for, the Poisson distribution is probably a good replacement for the ARMA-model. This distribution will only use positive and integer numbers. This distribution is given by

$$P(y = k) = \frac{e^{-\lambda}\lambda^k}{k!}, \tag{2}$$

where λ is the parameter which has to be estimated by Maximum Likelihood. This parameter should be a positive real number. This distribution is applicable for count data, because y and k are always integer-valued, which is also the case for count data. However, any autocorrelation cannot be covered by this distribution, so an extended version will be useful. These expansions are done in Section 2.3 and 2.4.

2.3 INAR(p)-Poisson model

The Integer-valued Autoregressive model is a method which is applicable to time series of count data, which uses the commonly known Poisson distribution.

The idea is that the data is decomposed in two parts: the first part will handle the autocorrelation in the data and the second part handles new arrivals. For example, a shop has some customers which will always go there, and probably some new customers which can not yet be covered by the autoregressive function. These new customers can than be added to the model by the new arrivals. This INAR-Poisson model is defined as

$$y_t = \alpha_1 \circ y_{t-1} + ... + \alpha_p \circ y_{t-p} + I_t, \tag{3}$$

where y_t is the observation at time t and the $\circ$ in this function is called the binomial thinning operator, which means that there is a chance of $\alpha_i \in [0,1]$ that a previous count will be transferred to the current period. This part can also be written as

$$\alpha \circ y_{t-p} \sim B(\alpha, y_{t-p}), \tag{4}$$

where $B(\alpha, y_{t-p})$ is the binomial distribution for y_{t-p} trials with probability of success α (Böckenholt, 1999, p. 321). From now on, p will be equal to 1, for simplicity.

I_t in (3) is a Poisson arrival process. These I_t are independently distributed Poisson variables with mean λ. This part can handle new information which was not observed yet, as in the example about customers given before.

Other variables can be included too. That can be done by replacing α or λ by a function of explanatory variables. Because $\alpha \in [0,1]$,

$$\alpha_t = \frac{e^{x_t'\gamma}}{1 + e^{x_t\gamma}} \tag{5}$$

can be used (Freeland and McCabe, 2002, p. 704). While λ can only be positive,

$$\gamma_t = e^{z_t'\beta} \tag{6}$$

seems to be a good replacement. In (5) and (6), x_t and z_t respectively are the possibly explanatory variables and γ and β are parameters.

This INAR-model seems applicable for count data, because it reckons with new arrivals which are not yet known from past information. Furthermore, it uses the fact that the variables are only positive and integer-valued. The INAR(1)-Poisson model assumes that the counts are Poisson distributed. This assumption seems better than the assumption of normality in the case of an ARMA-model. An other assumption of this method is that there is positive correlation between the observations, caused by the fact that $\alpha_i \in [0,1]$ in (3). The autocorrelation function

$$\rho_{t,t-1} = \alpha_t \sqrt{\frac{\lambda_{t-1}}{\lambda_t}}, \tag{7}$$

where λ is the positive Poisson parameter, shows that, according to the INAR-model, a high number of observations in y_{t-1} will lead to more observations in y_t and a low number in y_{t-1} will cause less observations in y_t (Böckenholt, 1999, p. 322). Probably, there is negative correlation in a certain count data set and than the INAR-model can not cover that. For example, one day a person buys a lot a bread. Than he will probably not buy that again the next day, because he still has enough. In this example, many counts yesterday can lead to less counts today. In this case, the INAR-model is probably less applicable.

All in all, the INAR-model seems to be pretty good for time series of count data: it reckons with the fact of positive and integer-valued variables.

2.4 ACI(p,q)-model

Another possibility of modeling time series is by using the newly suggested Autoregressive Conditional Integer-valued model. This method is inspired by the Autoregressive Conditional Duration model, introduced by Engle and Russell, 1998. Two things have to be accounted for: the distribution of the data and the parameter(s) which have to be calculated for this distribution. Because this article is about count data, the Poisson distribution with parameter λ is most valid, while this distribution makes use of integer-valued numbers. Again, other variables can be added by replacing the parameter for a formula, and so

$$\lambda_i = \omega + \alpha_1 y_{i-1} + ... + \alpha_p y_{i-p} + \beta_1 \lambda_{i-1} + ... + \beta_q \lambda_{i-q} \tag{8}$$

can be used (Engle and Russell, 1998, p. 1133). ω, α_i and β_i in this equation are parameters which can be estimated by using Maximum Likelihood and λ_i is the conditional expectation of y_i, $E(y_i|y_{i-1}, ...y_1)$, which can be used as the parameter in the Poisson distribution shown in (2).

Again, by choosing the numbers for p and q, one chooses how many lags should be used in the ACI-model. This paper will only use the ACI-model with p and q equal to 1. Of course, other distributions can be chosen, for example the Negative Binomial distribution. This choice depends on the data and which distribution describes this data best.

An asset of the ACI-model is that lags can be used, as shown in (8). Furthermore, by choosing the right distribution, the fact that only integer-valued variables are present can be taken into account. In contrast with the INAR-model proposed in Section 2.3, this method can handle negative correlations, by making α_i and/or β_i negative in the ACI-model in (8), which can be useful in the example given in Section 2.3. These properties are useful for the time series of count data, so probably this method can handle the count data pretty well.

To conclude, the ACI-model seems theoretically most applicable. This method has the least bounding assumptions which can harm the time series of count data. In the next section, simulations will be used to see which method is best in that particular view.

3 Model Selection by simulation

In Section 2 the different methods are shown and in this section, these methods will be compared by looking at a simulation. In this section simulations will be done to see if data generated by a certain method (such as a INAR- or ACI-model) can also be described by one of the other methods. This will show which method is applicable to different types of count data.

Simulation is a way to see which method is useable for all kind of data. First, in this section the simulation will be explained and after that, it will be shown which method can explain different simulated data in the best way, using different criteria, represented in Section 3.1.

In this paper, eight types of data, shown in Table 1, will be simulated: four types using the INAR(1)-model and four using the ACI(1,1)-model. In these types, some different simulations are included: with high or low autocorrelation and with many or little zeros. Than, the ARMA(1,1)-model, the INAR(1)-model, the ACI(1,1)-model and the Poisson distribution will be used to see which method is able to model the simulated data sets. In any further research, one can take a look at what happens when more lags are added. This is beyond the scope of this paper.

Some expectations can be made. In the first two simulations, the Poisson-model will do not so bad, because there is a low autoregressive part. In the fifth en sixth simulation, the INAR-model will not have the best outcome, because this method can not handle negative correlations. In the fourth and eighth simulation, the ARMA-model will stand a chance, due to a high autoregressive part and not so many zeros simulated in the data set. The presence of zeros indicates that the ARMA-model possibly will return some negative values, which is not applicable for count data.

	Simulation	Parameter values	% zeros	Autocorrelation
1	INAR	$\alpha \in [0.1, 0.25]$		
		$\lambda \in [1, 6]$	7.21	0.17
2	INAR	$\alpha \in [0.1, 0.25]$		
		$\lambda \in [7, 12]$	0.01	0.17
3	INAR	$\alpha \in [0.75, 0.9]$		
		$\lambda \in [1, 6]$	7.47	0.80
4	INAR	$\alpha \in [0.75, 0.9]$		
		$\lambda \in [7, 12]$	0.01	0.80
5	ACI	$\alpha \in [-0.2, -0.4]$		
		$\beta \in [-0.2, -0.4]$	8.14	-0.34
		$\lambda \in [1, 6]$		
6	ACI	$\alpha \in [-0.2, -0.4]$		
		$\beta \in [-0.2, -0.4]$	0.03	-0.34
		$\lambda \in [7, 12]$		
7	ACI	$\alpha \in [0.2, 0.4]$		
		$\beta \in [0.2, 0.4]$	8.03	0.33
		$\lambda \in [1, 6]$		
8	ACI	$\alpha \in [0.2, 0.4]$		
		$\beta \in [0.2, 0.4]$	0.03	0.33
		$\lambda \in [7, 12]$		

Table 1: Simulations

The simulation is done as follows (Ross, 2006, p. 49-56):

1. Generate a random variable U from the Uniform distribution between 0 and 1.

2. If $U < p_0$ set $y = 0$

3. If $U < p_0 + p_1$ set $y = 1$

4. If $U < p_0 + p_1 + p_2$ set $y = 2$
   ...

where p_i is the chance that y is equal to i, $P(y = i)$. For p_i different distributions can be chosen, for example the probability distribution of the INAR-model or the ACI-model. By doing this for a complete series of y, a time series data set is created, which can be used to see which method reflects this data best.

To see which model is best to use on a particular simulated data set, some criteria can be used. These criteria, which be explained in the next section, can be used to compare the methods for their ability to cover the data set.

3.1 Criteria

First, the Akaike and Schwartz Information Criteria (AIC and respectively SIC) are commonly used. By comparing two methods, the one with the lowest AIC or SIC is said to be the best. These criteria can be found in (9) and (10), where $l(\hat{\theta})$ is the log-likelihood value with optimal parameters, k the number of parameters used and n the number of observations (Franses and Paap, 2001, p. 42):

$$AIC = \frac{1}{n}\,(-2\;l(\hat{\theta}) + 2k) \tag{9}$$

and

$$SIC = \frac{1}{n}\,(-2\;l(\hat{\theta}) + k\;ln(n)). \tag{10}$$

Second, one can have a look at how good the methods can forecast the values of the simulated data set. These forecasts are done by looking at the expectations of the counted number y_{t+h}, which makes the Root Mean Squared Prediction Error (RMSE), defined by

$$RMSE = \sqrt{\frac{1}{m}\sum_{h=1}^{m}(y_{t+h} - \hat{y}_{t+h})^2} \tag{11}$$

appropriate to use. In this equation, m is the number of observations which have to be predicted and $\hat{y}$ is the forecast which is done. This forecast can be any real number. In this case, a part of the data is used to create the model, while the other part will be predicted using this model. Again, the model with the lowest value is said to be the best (Heij et al., 2004, p. 570).

Third, the power of prediction can be tested by using

$$\frac{1}{\sqrt{m}}(2B - m), \tag{12}$$

where m is the number of observations and B is a counted value of the times the prediction of model 1 differs less from the actual value than model 2 does. This equation is asymptotically $N(0,1)$ distributed. If this value is significantly smaller than 0, model 2 is better than model 1, and vice versa (Heij et al., 2004, p. 570).

Now, these criteria can be calculated every run of the simulation and than the average values can be compared to see which method is applicable to that data set.

3.2 Outcome of simulation

In *Appendix A*, the results of the simulation are shown. On the basis of this AIC, SIC and RMSE, a pecking order is made, which is represented in Table 2. Some interesting things can be noted.

Simulation	ARMA	INAR	Poisson	ACI
1	4	1	2	3
2	4	2	1	3
3	2	1	4	3
4	3	1	4	2
5	4	3	2	1
6	3	4	1	2
7	3	2	4	1
8	3	2	4	1
On average	3	1	4	2

Table 2: Pecking order

First of all, the simulation using an INAR-model can be covered by the Poisson-model pretty well, which can be seen by the value for the RMSE in the first simulation. But a problem with this method occurs when big autoregressive correlations are simulated, which can be seen by the big Akaike Information Criteria in the fourth simulation. So, the Poisson-model will not be safe to use when dealing with count data in time series. Moreover, it is obvious from *Appendix A* that the ARMA-model does not act very well, shown by high values of Akaike and Schwarz Criteria. Almost all of the times the ARMA-model is ranked third or fourth. Only when new arrivals are relatively low, as in simulation 3, the ARMA-model can cover the data. Furthermore, when the data is from the INAR-model, the ACI-model is moderate: most of the times it is ranked third in row.

Of course, the ACI-model is best when simulating data with this method. Furthermore, one can see that in simulations 7 and 8, INAR-model is not doing badly. Only when negative correlations are simulated, as happened in the fifth and sixth simulation, INAR is more worse than any other method. This indicates that the INAR-model is doing good, when simulation with the ACI-model is done with positive correlation.
Altogether, Table 2 shows that on average the INAR-model is doing best in explaining count data in time series.

It should be said that the criteria shown in *Appendix A* can not be tested very accurately. One can not say that one of the models is significantly better. That is why (12) is used too, and its values are shown in *Appendix B*. When the value in this *Appendix B* is greater than 1.96, method 1 is significantly better. When the value in this table is smaller than -1.96, method 1 is significantly worse. What can be seen from the first two simulations, is that the ACI-model does not significantly differ from the INAR-model. A problem occurs when a relatively high autoregressive part is added. Then the ACI-model does not predict the data very well.
When looking at the simulation using an ACI-model, one can see that the INAR-model is not working when the data is simulated with negative correlations. But, it is also clear, that this method does good when the autocorrelation is positive: it does not significantly differ from the ACI-model in the seventh and eighth simulations.

By looking at these simulations, it can be concluded that the ARMA-model is always doing badly when dealing with count data. The Poisson-model can not handle high autocorrelations, so this model seems not appropriate too. The ACI-model is not always capable of reproducing the data simulated by the INAR-model, especially in the presence of a high autoregressive part. The other way around seems to be much better: only when negative correlations are present, the INAR-model is significantly worse. And, as seen in *Appendix B*, in the last two simulations, the INAR-model does not significantly differ from the ACI-model. Also, in *Appendix A*, the INAR-model is on average the best method to use. That is why, according to simulations, the INAR-model seems best to use for count data in time series.

4 Application on Purchase Data

After simulated data, in this section real data will be used to see which model does best. The data that will be used in this paper is about the number of purchases of a certain product of a household every week. Obviously, the product which is observed will not always be bought, so this data set contains some zeros. Furthermore, different households will be looked at.

First, an initial analysis of the data can be done. In Table 3 some descriptive statistics of five of the observed households can be seen.

	Minimum	Maximum	Median	Mean	Variance
Household 1	0	4	0	0.294	0.732
Household 2	0	2	0	0.490	0.455
Household 3	0	3	1	0.863	0.921
Household 4	0	6	1	1.235	2.024
Household 5	0	8	2	1.588	3.167

Table 3: Descriptive Statistics

Most of the times, the product is not bought, which is shown by the low value of the median. The mean values found should be looked at in combination with the variance. This variance is in four out of five cases even higher than the mean, which is called overdispersion. Furthermore, there seems to be correlation between lagged observations, so probably an autoregressive model is recommended.

In *Appendix C* the Information Criteria, AIC and SIC, are indicted. Because not so many weeks are available, the decision is made not to use part of the data for prediction. That is why it is not possible to use the Root Mean Squared Prediction Error (RMSE). The first thing to note is that none of the methods is good for this data set. That is because no other explanatory variables are added, which could have made the methods more appropriate for the data. Not surprisingly, the ARMA-model is doing worst, caused by the high number of zeros in the purchase data. Furthermore, what we can see is that in all cases, the Poisson-model is doing pretty good. In this model, no lagged variables are added, so this indicates that these lagged observations do not have very significant influence. But, it is also clear that the ACI-model does well too, especially when autocorrelation seems to be relatively high. In the Poisson-model, for every point in data the same parameter λ, shown in Table 4, is used and this can be seen as the expected value of the number of purchases.

Household	1	2	3	4	5
Parameter value	0.24	0.50	0.88	1.26	1.58

Table 4: Poisson parameters

Using the ACI-model from (8), the parameters in Table 5 are found.

Household	1	2	3	4	5
ω	0.084	0.118	0.411	0.111	0.849
α	0.116	0.257	0.121	0.357	-0.175
β	0.493	0.528	0.421	1.220	0.630

Table 5: ACI parameters

It can be concluded that when the purchase data are not really autoregressive, the Poisson-model should be used and when more autocorrelation is found, the ACI-model is a good replacement.

5 Conclusion

After having shown the assumptions in Section 2, the simulation in Section 3 and the purchase data in Section 4 three conclusions can be drawn: a theoretical one, a conclusion based on simulations and a practical conclusion based on household purchases.

Theoretically, the Autoregressive Condition Integer-valued model is best for count data in time series. This model accounts for new arrivals, autocorrelation and for negative as well as positive correlations.
Based on simulations, the Integer-valued Autoregressive model does pretty well. This model can even handle data which is simulated by another method, such as the ACI-model. The only problem is that negative correlations will not be covered properly by the INAR-model.
Practicallly, for purchase data used in this paper, the Poisson-model and ACI-model do best. This is mostly due to the fact that no lagged variables are really significant to use, which makes the ARMA-model and the INAR-model less appropriate. When there is small autocorrelation, the ACI-model covers this best.

All in all, because the INAR-model is in almost all simulations acceptable, this model seems best to use in time series of count data. The ACI-model is good, but when autocorrelation is high, the INAR-model is significantly better than this method. Only when it is clear that negative correlations will occur, the ACI-model is better to use.

6 Further Research

Some additional things can be watched at in further research. First, as already mentioned in Section 2, the methods can be extended by adding other (possibly) explanatory variables. This will make the methods better in predicting than they are now, in this paper. Second, maybe a Negative Binomial Distribution will do better in this kind of data than the Poisson Distribution does. This can be investigated too. Third, the methods can be extended by adding more lags. This will probably return different answers, but for the INAR(p)-model, this is not a simple task. Moreover, other data, with a bigger autoregressive part, will give another answer than is given in this paper. These are just four possible further items which can be looked at. Probably more research can be thought of.

7 References

Böckenholt, U. (1999): Mixed INAR(1) Poisson regression models: Analyzing heterogeneity and serial dependencies in longitudinal count data, Journal of Econometrics 89.

Engle, R.F. and Russell, J.R. (1998): Autoregressive conditional duration: A new model for irregularly spaced transaction data, Econometrica, Vol. 66, No. 5.

Franses, P.H. and Dijk, D. van (2009): Time Series Models for Business and Economic Forecasting.

Franses, P.H. and Paap, R. (2001): Quantitative models in marketing research.

Freeland, R.K. and McCabe, B.P.M. (2002): Analysis of low count time series data by Poisson autoregression, Journal of time series analysis, Vol. 25, No. 5.

Heij, C., Boer, P. de, Franses, P.H., Kloek, T. en Dijk, H.K. van (2004): Econometric Methods with Applications in Business and Economics (1st ed.).

Quddus, M.A. (2008): Time series count data models: An empirical application to traffic accidents, Accident analysis and prevention 40.

Ross, S.M. (2006): Simulation (4th ed.).

A Simulation: Information Criteria

	Simulation	Parameter values	Criteria	ARMA	INAR	Poisson	ACI
1	INAR	$\alpha \in [0.1, 0.25]$	Akaike	5.27	3.87	3.88	3.87
		$\lambda \in [1, 6]$	Schwarz	5.33	3.91	3.90	3.93
			RMSE	1.836	1.824	1.818	1.826
2	INAR	$\alpha \in [0.1, 0.25]$	Akaike	6.67	5.00	4.98	5.00
		$\lambda \in [7, 12]$	Schwarz	6.73	5.04	4.99	5.06
			RMSE	3.054	3.068	3.043	3.076
3	INAR	$\alpha \in [0.75, 0.9]$	Akaike	3.75	2.72	3.81	3.23
		$\lambda \in [1, 6]$	Schwarz	3.81	2.76	3.83	3.29
			RMSE	1.685	1.787	1.823	1.805
4	INAR	$\alpha \in [0.75, 0.9]$	Akaike	5.23	3.87	4.96	4.38
		$\lambda \in [7, 12]$	Schwarz	5.29	3.91	4.98	4.44
			RMSE	2.879	3.037	3.099	3.045
5	ACI	$\alpha \in [-0.2, -0.4]$	Akaike	5.37	4.07	4.05	3.90
		$\beta \in [-0.2, -0.4]$	Schwarz	5.43	4.12	4.07	3.96
		$\lambda \in [1, 6]$	RMSE	1.997	1.988	1.987	1.988
6	ACI	$\alpha \in [-0.2, -0.4]$	Akaike	6.76	5.20	5.16	5.03
		$\beta \in [-0.2, -0.4]$	Schwarz	6.82	5.24	5.18	5.09
		$\lambda \in [7, 12]$	RMSE	3.277	3.282	3.271	3.279
7	ACI	$\alpha \in [0.2, 0.4]$	Akaike	5.34	3.89	4.00	3.88
		$\beta \in [0.2, 0.4]$	Schwarz	5.40	3.93	4.02	3.94
		$\lambda \in [1, 6]$	RMSE	1.966	1.978	1.980	1.979
8	ACI	$\alpha \in [0.2, 0.4]$	Akaike	6.78	5.05	5.16	5.04
		$\beta \in [0.2, 0.4]$	Schwarz	6.84	5.09	5.18	5.10
		$\lambda \in [7, 12]$	RMSE	3.268	3.287	3.289	3.287
	On average		Akaike	5.65	4.21	4.50	4.29
			Schwarz	5.71	4.25	4.52	4.35
			RMSE	2.495	2.283	2.539	2.536

Information Criteria per simulation: the lower, the better it is.

B Comparing methods

Method 1	ARMA	ARMA	ARMA	INAR	INAR	Poisson
Method 2	INAR	Poisson	ACI	Poisson	ACI	ACI
Simulation 1	-18.300	-17.969	-18.220	2.343	0.501	1.512
Simulation 2	-16.824	-16.672	-17.280	2.835	0.590	0.787
Simulation 3	-59.847	-63.808	-60.508	8.953	3.515	-3.640
Simulation 4	-55.446	-61.733	-60.052	5.510	3.801	-5.769
Simulation 5	-12.02	-11.99	-13.81	-11.42	-2.80	-2.56
Simulation 6	-12.44	-12.47	-11.19	-9.56	-2.39	-2.42
Simulation 7	-26.52	-26.81	-28.14	2.58	-1.47	-0.82
Simulation 8	-26.79	-27.45	-27.30	2.61	-0.95	-1.12

Comparing methods by using (12). When significantly positive, model 1 is better, when significantly negative, model 2 is better.

C Data: Information Criteria

	ARMA	INAR	Poisson	ACI
Household 1				
Akaike	2.60	1.43	1.41	1.44
Schwarz	2.72	1.51	1.45	1.56
Household 2				
Akaike	2.04	1.83	1.84	1.82
Schwarz	2.15	1.91	1.87	1.93
Household 3				
Akaike	2.83	2.53	2.50	2.55
Schwarz	2.94	2.60	2.53	2.66
Household 4				
Akaike	3.61	3.25	3.21	2.87
Schwarz	3.73	3.33	3.25	2.98
Household 5				
Akaike	4.04	3.81	3.77	3.78
Schwarz	4.16	3.89	3.81	3.89

Information Criteria per household: the lower, the better it is.