

DEPARTMENT OF MARKETING

Forecasting Newspaper Demand with Censored Regression

Meltem Kiygi Calli & Marcel Weverbergh

UNIVERSITY OF ANTWERP
Faculty of Applied Economics



Stadscampus
Prinsstraat 13, B.213
BE-2000 Antwerpen
Tel. +32 (0)3 220 40 32
Fax +32 (0)3 220 47 99
<http://www.ua.ac.be/tew>

FACULTY OF APPLIED ECONOMICS

DEPARTMENT OF MARKETING

Forecasting Newspaper Demand with Censored Regression

Meltem Kiygi Calli & Marcel Weverbergh

RESEARCH PAPER 2008-006
APRIL 2008

University of Antwerp, City Campus, Prinsstraat 13, B-2000 Antwerp, Belgium
Research Administration – room B.213
phone: (32) 3 220 40 32
fax: (32) 3 220 47 99
e-mail: joeri.nys@ua.ac.be

The papers can be also found at our website:
www.ua.ac.be/tew
(research > working papers)

D/2008/1169/006

Forecasting Newspaper Demand with Censored Regression

Meltem Kiygi Calli

University of Antwerp, 2000 Antwerp, Belgium, meltem.kiygicalli@ua.ac.be

Marcel Weverbergh

University of Antwerp, 2000 Antwerp, Belgium, marcel.weverbergh@ua.ac.be

Abstract

Newspaper circulation has to be determined at the level of the individual retail outlets for each of the editions to be sold through such outlets. Traditional forecasting methods provide no insight into the impact of the service level defined as the probability that no out-of-stock will occur. The service level results in out-of stock situations, causing missed sales and oversupply or returns. In our application management sets a policy aiming at a 97 percent service level. The forecasting system developed provides estimates for excess deliveries and for the expected shortages. The results compare favorably to the traditional moving average approach previously employed by the publisher. Censored regression is a natural approach to the newspaper problem. It provides information on key policy variables and it is relatively simple to integrate into the distribution policy, with only small adaptations to the existing forecasting and distribution policy.

Keywords: Censored Regression, decision analysis, forecasting, distribution

Introduction

Newspaper distribution is similar to the distribution of highly perishable goods with a shelf life of one day. The problem is different from the distribution of other perishable goods in the sense that supply (the number of copies printed and distributed) is under full control of the publisher, as opposed to other categories of perishable goods and services, where supply is exogenous, like for fresh produce, or fixed as in the case of theatres and amusement parks. This makes it necessary to develop a short-term forecasting system and a distribution system which takes into account the interaction between the service strategy and sales. The total circulation of a newspaper is based on two components: the subscriptions, which are not part of our investigation, since they do not generate daily random fluctuations in sales and free float in newsstands and newspaper shops. The number of newspapers delivered is adjusted for each point of sales on a daily basis, in order to achieve the desired service level, with a minimum number of returns. The company makes a distinction between normal days, for which an automatic forecasting and delivery system is applied based on a moving average, and special days, for which the circulation manager makes an adjustment to the normal deliveries by means of a judgmental modification of the number of newspapers delivered. Such special days may be caused by news events, marketing actions, local events and so on.

The moving average forecasting system does take into account the censoring effect of the distribution policy: on days that an out of stock occurs the potential sales that could have been realized are higher than the observed sales: the reaction is to increase deliveries in order to avoid future shortages. However, the amount of this increase is largely judgmental. The use of censored regression models helps to devise a policy with a better statistical underpinning. A censored regression allows a formal assessment of the impact of a specific service policy in terms of service level and the trade-off with respect to excess supply it generates. At the same time it allows a more formal identification of the influences which are relevant in the forecasting system.

Censored regression analysis goes back to the seminal article of Tobin (1958), in which automobile purchases are analyzed. From these origins censored regression is also called a Tobit model. In linear regression models ordinary least squares estimators are unbiased and consistent even when the assumption of normality of the disturbances is violated. However, under censoring this is no longer true, which leads to the need of appropriate estimation

techniques. Heckman (1979) develops a technique which corrects for a selection bias due to censoring in linear regression models with normal errors. This technique revealed that likelihood-based estimators are consistent and normally distributed. Dubin and Rivers (1989) consider Heckman's (1979) version of the Tobit model. They demonstrate how the Heckman model can be applied to probit or logit models. Miller (1976) revealed a new estimator which is called Kaplan-Meier Least Squares and it is derived by minimizing the weighted sum of squares of the residuals. Buckley and James (1979) discuss a parameter estimation method in the linear regression model in which the dependent variable is censored. In Buckley and James' (1979) simulation study, estimates of 8 different cases are compared with the estimates of Miller (1976). According to the simulation results, Buckley and James' estimates are almost unbiased in all 8 cases despite the changing censor patterns while Miller's estimates are biased if censoring is unevenly spread over time. Normally distributed error terms have a distribution functions with a known parametric form. Estimation methods by Miller (1976) and Buckley and James (1979) propose a nonparametric estimation of error term distribution. Koul *et al* (1981) and Powel (1986) also propose new estimators in a linear regression model in which the dependent variable is censored. Koul *et al*'s estimator is applicable when the observations are randomly censored on the right and when the error distribution is unknown. Powel (1986) proposes a censored 'Least Absolute Deviations' (LAD) estimator which is consistent and asymptotically normally distributed. Lu and Burke (2005) propose a technique called 'Censored Average Derivative Estimation' (CADE) for nonparametric censored regression models with randomly censored samples. This technique involves three steps: (1) transformation of the censored data into synthetic data or pseudo-responses, (2) estimation of the average derivative estimators of the regression function, (3) approximation of the unknown regression function by means of a univariate regression using a nonparametric censored regression technique. In sum, in most of the models, in order to obtain the estimates of the censored regression model, either a parametric form or a parametric error distribution is assumed, the average derivatives estimation technique providing an alternative to these approaches. The robustness of the Tobit estimator with respect to heteroskedasticity and to non-normality has been considered by Arabmazar and Schmidt (1981 and 1982). According to these studies, heteroskedasticity of the disturbance can cause inconsistency of the parameter estimates and the bias resulting from non-normality can be substantial. The bias resulting from non-normality is more important when the disturbance variance is unknown. Robustness of the estimators in the case of heteroskedasticity is also discussed in Schneider (1986) on the basis of simulation results.

Censored regression models have been discussed extensively (Greene, 1997; Jöreskog 2002; Schneider 1986). They have been applied in a variety of applications (Tobin, 1958; Heskes *et al* 2003). The Cox proportional hazards model is widely used in the analysis of censored time-to-event data. Hamasaki and Kim (2007) consider a Box and Cox power-transformation for non-normal censored regression. A simulation study is performed for several simple regression models and allows concluding that reasonable estimates of parameters can be obtained.

Heskes *et al* (2003) consider a software system, which is called ‘just enough delivery (JED)’ for optimizing newspaper sales. This system is based on a combination of neural and Bayesian methodology. This Bayesian inference tries to take into account the similarities between different outlets. Because of non-persistent random shocks (outlet sensitivity to weather circumstances or seasonal effects, special events, marketing actions), outlet sales are hard to predict. In their application, a correction related to nonstationarity is added to the forecast. The addition of this correction term to the prediction model makes the system very robust against unpredictable changes. This software system has been tested for several newspapers and magazines.

In this paper, we discuss the application of censored regression, with the objective of developing a delivery policy which achieves a given service level. The structure of the paper is as follows. First we situate the decision problem in relation to the service level decision. Next we discuss the data, model specification and estimation results, followed by an evaluation of the delivery policy implications. In the last section we discuss the implementation and the main conclusions of our analysis.

The decision problem

A decision model for the optimal number of copies of a newspaper to be delivered at a particular point of sales is straightforward, and can be sketched in its simplest form as follows:

Assume the probability density function $f(S)$ of the number of newspapers to be sold at a particular outlet on a given day is known, and that the marginal revenue per unit sold, R , is

constant. Denote C as the cost per unit of unsold copies. Then, defining L as the number of copies provided the payoff function is:

$$\pi = R \left(\int_0^L x f(x) dx + L \int_L^{\infty} f(x) dx \right) - C \int_0^L (L-x) f(x) dx \quad (1)$$

Optimizing with respect to L leads to the following optimality condition:

$$F[L^*] = \int_0^{L^*} f(x) dx = \frac{R}{R+C} \quad (2)$$

The optimal number L^* to be delivered is such that the service level, this means the probability of meeting demand, equals the ratio of $R/(R+C)$. The company aims for a service level of 0.97. If optimal this implies $R = C F(L^*) / [1 - F(L^*)] = 32.33 C$, which means that a service level of 0.97 assumes a ratio of marginal revenue to marginal cost of approximately 32. One conclusion related to the optimal delivery policy is that it should aim at the same service level for all outlets, irrespective of the size or the variability, as long as all outlets are characterized by the same marginal revenue and cost. However differences in variability do imply differences with respect to the expected shortages and returns.

The expected number of returns and the expected number out of stock are crucial in the evaluation of the service strategy. This implies that non-parametric methods are less indicated here. We will assume that a normal distribution applies for the demand for newspapers. In this case the relevant magnitudes can be easily computed (Green, 1997; Jöreskog, 2002).

Assume the following model for unit sales at the level of the point of sales/edition, where S_c is number of copies sold (censored), S is (uncensored) demand, and L is the number delivered. Assume that S is normally distributed with mean μ and standard deviation σ . Censored sales are equal to S if $L \geq S$ and equal to L otherwise. The bias in the estimates of the parameters resulting from censoring are discussed in Greene (1997) and Jöreskog (2002). Here the expected number of missed sales (M) and the expected number of unsold copies (U) at a given service level are the central concerns. Missed sales are the difference between the uncensored and the censored mean: $M = \mu - \mu_c$, while the expected number of returns U is $L - \mu_c$. For example, if S is $N[60, 8]$ and $F[L]=0.97$, then the required number of copies L^* is $\mu + 1.88\sigma = 75$, resulting in $\mu_c = 59.90$, $M = \mu - \mu_c = 0.094$, while the unsold copies U are

equal to $L - \mu_c = 15.14$. This example yields a return rate of 20.2 percent of the copies delivered, not untypical for the daily newspapers business.

---Insert Figure 1 about here---

Figure 1 displays the relationship between the service level chosen and the number of returns for two examples, $S = N[60,8]$ and $S = N[60,16]$. The figure shows that a higher variability increases the average number of returns associated with the service level if the latter is higher than 0.5, and this effect becomes more severe as the service level increases. At the same time a higher variability is associated with an increase in missed sales, although this effect is much less pronounced.

The data

The available data cover the period from January, 3 2005 to April, 1 2006. They relate to 1321 unique SKU's or stock keeping units, which refers to different editions of the newspaper and/or outlets. This results in a total of 491602 data points. The number of observations per SKU varies, and depends for instance on holiday breaks in a given retail outlet, on weekly closing days in some outlets and possibly on occasional, irregular closing dates. The maximum number is about 360 days for one SKU. When there were less than 250 days the SKU was deleted from further analysis. Also, recorded days where sales were abnormally low, indicating that the shop was effectively closed were eliminated. If a particular weekday had less than 10 active occurrences for a particular outlet it was deleted from the analysis also. This data cleaning process resulted in a total of 1274 SKU's selected for analysis, with 454935 observations.

The average number of daily deliveries is 80730 or 63.37 per SKU/day. The average number of unsold copies or returns is 16321 per day. This means that the observed return rate is 20.22 percent. A count of the number of SKU-days with out of stock shows an average service level of 93 percent, 4 percent below the stated target of 97 percent. This is an unweighted service level, the weighted service level, which takes into account the average number of copies sold per SKU, is 94 percent indicating that a somewhat higher service level is obtained for larger SKU's than for smaller ones.

The data contain the number of copies sold per day (S), the number of copies delivered (L), the extra copies provided, this means the number of copies delivered in excess of normal supplies, an indicator if management deemed such an increase in deliveries necessary (publisher initiated increase, denoted D^P) and an indicator whether the retailer requested an increase in the number of copies (retailer initiated increase denoted D^R). Such situations occur relatively frequently. Especially days for which D^P applies, which occurs on 25% of the days, result in substantial increases in deliveries. Requests from retailers are observed on 17 percent of the days. The latter do not always lead to actual increases, but we include them in the models in order to test the degree to which sales actually show increases on such days.

For calculating the normal number of newspapers to be delivered for each SKU/day combination, the publisher uses a weighted average as follows:

$$L_{id,d} = \sum_{i=1}^5 w_i S_{id,d-7i} + z \sigma_{id,d} \quad (3)$$

The parameter z has to be determined on the basis of the required service level. The weighted average is applied for each weekday and SKU separately, and runs over the last 5 observations on normal days; this means it is based only on days that no special conditions or extra deliveries are in effect. When a special occasion takes place, the circulation manager increases the number of copies delivered on a judgmental basis.

Specification and estimation results

A preliminary analysis of the data shows low to no seasonality and fair to good stationarity for most SKU's. Also autocorrelations are generally weak. A censored regression is applied for each of the SKU's with the specification given in equation (4)

$$\begin{aligned} S_{id,i}^c &= L_{id,i} && \text{if } S_{id,i} \geq L_{id,i} \\ &= \beta_0 + \sum_{d=1}^5 \beta_{id,d} \text{Day}_{d,i} + \beta_{id,6} D_{id,i}^R + \beta_{id,7} D_{id,i-6}^R && (4) \\ &+ \beta_{id,8} D_{id,i}^P + \beta_{id,9} D_{id,i-6}^P + \beta_{id,10} S_{id,i-6}^c + \varepsilon_{id,i} && \text{if } S_{id,i} < L_{id,i} \end{aligned}$$

This specification assumes that the data from a particular SKU can be pooled over the weekdays, but are accounted for by means of dummy variables.

The explanatory variables are 5 daily dummies (there are no Sunday editions), a dummy for days where the number of deliveries is increased (D^P), which indicates that management expects increased sales, a dummy for deviations in D^R , for days that the retailer expects increased sales (irrespective of whether this request was implemented or not), and observed sales one week earlier. We also tested the inclusion of one day lag of sales. This yields a significant coefficient for 80 percent of the cases, but it was decided to drop it because information on sales the day before is not available in time and therefore is not relevant for forecasting. Further the model incorporates D^R and D^P one week lagged, in order to moderate the impact of sales the week before, which might be inflated due to special circumstances.

The explanatory power of the variables varies greatly over the sample. This is obvious from the range observed for the coefficient of determination, which is between 0.45 and 0.85. The one week lag for extra deliveries, which should correct for higher than normal sales a week before, without necessarily a carry-over into the next week, is significantly different from zero at the 95% confidence level for 22 percent of the SKU's. The coefficient D^P is significant in 52 percent of the cases. Retailers requests D^R are significant for 50 percent of the cases, the one week lag of D^R in only 10 percent. The intercept measures expected sales on Saturdays; sales on other weekdays are mostly significantly lower. Sales one week lagged are significant in most cases (see Table 1).

---Insert Table 1 about here---

Validation

The validation tests are of crucial importance in this study, since the service level strategy is directly related to the normal probability distribution assumed for the residuals in equation (4). Also the assumption of homoskedasticity is crucial to the service level strategy, as is obvious from the example in Figure 1. Both assumptions have a direct impact on the accuracy of the delivery policy.

Normality

The hypothesis of normality of the residuals in equation (4) has to be rejected in the majority of the cases (57% Kolmogorov-Smirnov, 65% Cramer-von Mises, 71% Anderson-Darling and 84% for Shapiro-Wilk tests). It is recommended in such cases to apply a logarithmic model or to apply Box-Cox transformations, the logarithmic and linear models being at opposite extremes of the class of Box-Cox models. However, applying the logarithmic model did not lead to substantial improvement as far as normality tests were concerned. We discuss the impact of non-normality in the section about the service level strategy.

Heteroskedasticity

Also the tests of heteroskedasticity fail in an important number of cases. We test this by the following variant of the White test (White, 1980).

$$\hat{\varepsilon}_{id,i}^2 = \sum_{d=1}^5 \alpha_{id,d} D_{id,i} + u_{id,i} \quad (5)$$

---Insert Figure 2 about here---

All weekdays yield a significant coefficient in approximately the same number of cases (35 to 40 percent, see Figure 2). This result has a direct impact on the service strategy. Figure 1 and the discussion about the optimal number of newspapers delivered to a particular outlet demonstrated that a higher variability of sales increases the deliveries necessary to obtain the service level required. Therefore, in the presence of heteroskedasticity, days with higher variability will be undersupplied, while days with lower variability will be oversupplied.

In order to resolve this problem the model in equation (4) is modified and re-estimated as follows:

$$\begin{aligned} S_{id,d,i}^c &= L_{id,d,i} && \text{if } S_{id,d,i} \geq L_{id,d,i} \\ &= \beta_{id,d,0} + \beta_{id,d,6} D_{id,d,i}^R + \beta_{id,d,7} D_{id,d,i-6}^R && \\ &+ \beta_{id,d,8} D_{id,d,i}^P + \beta_{id,d,9} D_{id,d,i-6}^P + \beta_{id,d,10} S_{id,d,i-6}^c + \varepsilon_{id,d,i} && \text{if } S_{id,d,i} < L_{id,d,i} \end{aligned} \quad (6)$$

This means a separate censored regression is estimated for each combination of weekday and pos/SKU. The predictive power of these regressions is limited, D^R and D^P are significant in 29 and 27 percent of the cases, and only the $S_{id,d,i-6}$ reaches the 95% significance level in 45 percent of the cases (see Table 2 and Figure 3), lending support to the overall conclusion that the process is fairly stable on a day by day basis for most of the SKU's covered in the analysis. The coefficient of determination varies greatly between 0.45 and 0.85 (see Figure 4).

---Insert Table 2 about here---

---Insert Figure 3 about here---

---Insert Figure 4 about here---

Normality is much better respected for equation (6) than for equation (4): the Kolmogorov-Smirnov test fails only in 22% of the cases, (Cramer-van Mises 26%, Anderson-Darling 29%, and Shapiro-Wilk 35%). It appears that non-normality is related to the presence of outliers in both tails of the distributions, and a somewhat higher kurtosis than what is to be expected for the normal distribution. The probability plot for one SKU, typical of deviations from normality is shown in Figure 5. We find a small impact of the non-normality on the service level, as will be discussed below.

---Insert Figure 5 about here---

Delivery policy evaluation

A delivery policy based on the results of the censored regression is simple. The expected number of newspapers sold, $\hat{S}_{id,d}$ is derived from the estimation results of equation (6). The required deliveries are $L_{0.97} = \hat{S}_{id,d} + 1.881\hat{\sigma}_{id,d}$. The estimated standard deviation of sales depends slightly on the values of the explanatory variables, but as a close approximation it can be taken equal to the standard error of the residuals in equation (6).

Evaluation of the service policy is done on the basis of missed sales and of returns. We compute the deliveries required under a 97 percent service strategy based on both models. This allows making the following comparisons:

- A backtracking evaluation, providing insight into the accuracy of the two models with respect to service level and the hypothetical shortages and returns it would have caused if applied in the past. It is approximate because on days where the sales are censored the true level of sales is not observed which may result in downward bias of the estimated shortage.
- The theoretical shortage obtained from equations (4) and (6)

---Insert Table 3 about here---

Tables 3 presents the results obtained by backtracking compared to the deliveries, shortages and returns for two service levels: the 97% service level aspired to and the 93% service level realized in the sample period. Equation (4) yields an apparent service level of 97.8%, while equation (6) gives an apparent service level of 97.9%. So a small overachievement of the target service level of 97% is observed. Daily deliveries of 82796 for equation (4) and 81623 for equation (6) are required, to be compared to an observed average of 80730. Increasing the service level therefore leads to an increase in deliveries of approximately 900 copies per day compared to current practice. Equation (4) gives backtracking shortages of 192 copies per day for a 97% strategy and of 326 for a 93% strategy. The results for equation (6) are slightly better with a backtracking shortage of 154 and of 284 respectively. Taking into account heteroskedasticity therefore reduces missed sales with 38 copies per day.. The backtracking shortages for a 97 percent service strategy are higher than those theoretically obtained. Equation (4) predicts a shortage of 111 per day, which is 81 less than the backtracking results, while equation (6) gives a result of 104 copies, close to the backtracking result. For a 93% service strategy equations (4) and (6) yield theoretical shortages of 296 versus 276 per day respectively.

The superiority of equation (6) over equation (4) is somewhat more pronounced in terms of returns. A sample average of 16321 or 12.81 per SKU/day excess copies is observed per day which is approximately a 20% return rate. Equation (4) at a 97% service level requires higher deliveries and also higher returns of 18579 per day which amounts to 22% of deliveries. Equation (6) performs better, with 17368 returns per day (21%). At a 93% service level the returns drop considerably to 14839 per day which is 19% of deliveries for equation (4) and 13883 per day for equation (6) which amounts to 18%.

The theoretical results give lower returns than the backtracking results, 18092 for equation (4) per day and 16894 for equation (6) per day at a 97% service level. For the 93% service level, equation (4) and equation (6) yield theoretical returns of 14405 per day and 13451 per day respectively.

---Insert Table 4 about here---

On days where both D^R and D^P are zero (normal days), the observed deliveries are 63.6 per SKU/day on average (see Table 4). Equation (6) requires a somewhat higher average of 64.3. On days where D^R is active average deliveries are close to the average required. For Days with D^R flag active, the observed deliveries exceed the 97% service level supplies by 5.7% on average. This shows that the increase in deliveries required for the desired service level on normal days is compensated to some extent by reducing the deliveries on special days.

The results show that non-normality has a noticeable but limited impact on the results. The deviation from normality is linked to a slightly higher kurtosis than would be expected under normality. The backtracking shows that the service level actually achieved exceeds the target slightly. At the same time the theoretical shortage is lower than found in the backtracking evaluation. This is linked to the presence of larger outliers than would be expected, indicating that the extreme sales levels, when they occur, tend to be larger than in a normal distribution (see Figure 6).

---Insert Figure 6 about here---

Implementation

Technically implementation of the results is not a difficult issue. The estimation results can be linked to the forecasting system, but of course some practical issues have to be solved, and the forecasting and delivery system has to be adapted for effective results. The two main issues are 'exception handling', this means selection of outlets and days for which a judgmental intervention is required (recently established outlets, days or editions not covered by the system). Further, also related to exception handling is the policy of extra-deliveries which according to the analysis, tend to be excessive, leading to lower service levels on normal days but larger than average returns on special days. The publisher needs to weigh the importance

of marketing objectives on such days with efficiency considerations. The second issue related to implementation is the frequency of recalibration. While the results show sufficient stability, the system is not 'self-monitoring' in the sense a moving average forecasting system would be. A yearly estimation cycle, with monitoring of outlets requiring judgmental assessment seems to provide an adequate balance of such considerations.

Conclusions

The present paper has investigated the use of censored regression and the possibilities it offers to answer the core problems associated with defining a service strategy for perishable goods. These relate primarily to the evaluation of the service strategy in terms of out of stock and in term of excess supply. The application provides two qualitative messages of importance for management: increasing the service level results in very low increases in sales, the expected returns are much more affected by high service levels.

Moderate deviations from normality are observed in our data. Since the optimal delivery policy is closely linked to right tail of the sales distribution, this needs attention. We demonstrate that the impact on the delivery policy and key policy variable is minimal however. There is a potential underestimation of missed sales, linked to a small overestimation of the service level because of a somewhat higher kurtosis than expected under normal distribution theory. The lack of normality is less damaging than the presence of heteroskedasticity, which if not treated adequately will result in a deterioration of the delivery policy such that higher deliveries lead to higher excess supplies, without a marked effect on sales.

Our analysis is subject to some potential limitations. Censored regression allows a straightforward evaluation of the service strategy as long as the time series which is analyzed is reasonably stationary, and provided that the structure of the model is sufficiently similar for all retail outlets under analysis. However, the potential merits of automated stepwise procedures, where the model structure is directly modified for each outlet remains to be seen.

Also procedures to screen and adjust the model for non-stationary time periods would be beneficial. Last the problem of updating of the estimates and the optimal length of the time period remains unexplored. One approach could be to develop adaptive procedures which

combine smoothing or weighted average forecasting procedures with censored estimation. This analysis demonstrates however the practical relevance even with the constraints of currently available software.

The primary objective of the study was to estimate the lost sales incurred because of out of stock situations. The study shows this impact is quite limited at the service levels considered, and that the trade off against excess supplies is quite substantial.

References

- Arabmazar, A., Schmidt, P., (1981). Further Evidence on the Robustness of the Tobit Estimator to Heteroskedasticity. *Journal of Econometrics* **17**: 253-258.
- Arabmazar, A., Schmidt, P., (1982). An Investigation of the Robustness of the Tobit Estimator to Non-Normality. *Econometrica* **50**, 4: 1055-1063.
- Buckley, J., James, I., (1979). Linear Regression with Censored Data. *Biometrika* **66**, 3: 429-436.
- Dubin, J. A., Rivers, D., (1989). Selection Bias in Linear Regression, Logit and Probit Models. *Sociological Methods and Research* **18**, 2&3: 360-390.
- Greene, W.H., (1997). *Econometric Analysis*. Prentice-Hall. Inc.: New Jersey.
- Hamasaki, T., Kim, S.Y., (2007). Box and Cox Power-Transformation to Confined and Censored Nonnormal Responses in Regression. *Computational Statistics and Data Analysis*. Doi: 10.1016/j.csda.2006.12.015.
- Heckman, J.J., (1979). Sample Selection Bias as a Specification Error. *Econometrica* **47**, 1: 153-162.
- Heskes, T., Spanjers, Jan, Spanjers, Joost, Bakker, B., Wiegerinck, W., (2003). Optimising Newspaper Sales Using Neural-Bayesian Technology. *Neural Computing and Applications*. 12:212-219, Doi: 10.1007/s00521-003-0384-x.
- Jöreskog, K.G., (2002). Censored Variables and Censored Regression. <http://www.ssicentral.com/lisrel/techdocs/censor.pdf>
- Koul, H., Susarla, V., Van Ryzin, J., (1981). Regression Analysis with Randomly Right-Censored Data. *The Annals of Statistics* **9**, 6: 1276-1288.
- Lu, X., Burke, M.D., (2005). Censored Multiple Regression by Method of Average Derivatives. *Journal of Multivariate Analysis* **95**: 182-205.
- Miller, R.G., (1976) Least Squares Regression with Censored Data. *Biometrika* **63**: 449-464.

- Powell, J.L., (1986). Censored Regression Quantiles. *Journal of Econometrics* **32**, 1: 143-155.
- Schneider, H., (1986). *Truncated and Censored Samples from Normal Populations*. Marcel Dekker, Inc.: New York.
- Tobin, J., (1958). Estimation of Relationships for Limited Dependent Variables. *Econometrica* **26**: 24-36.
- White, H., (1980). A heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. *Econometrica* **48**, 4: 817-838.

Figure 1 Expected Returns as a Function of Service Level

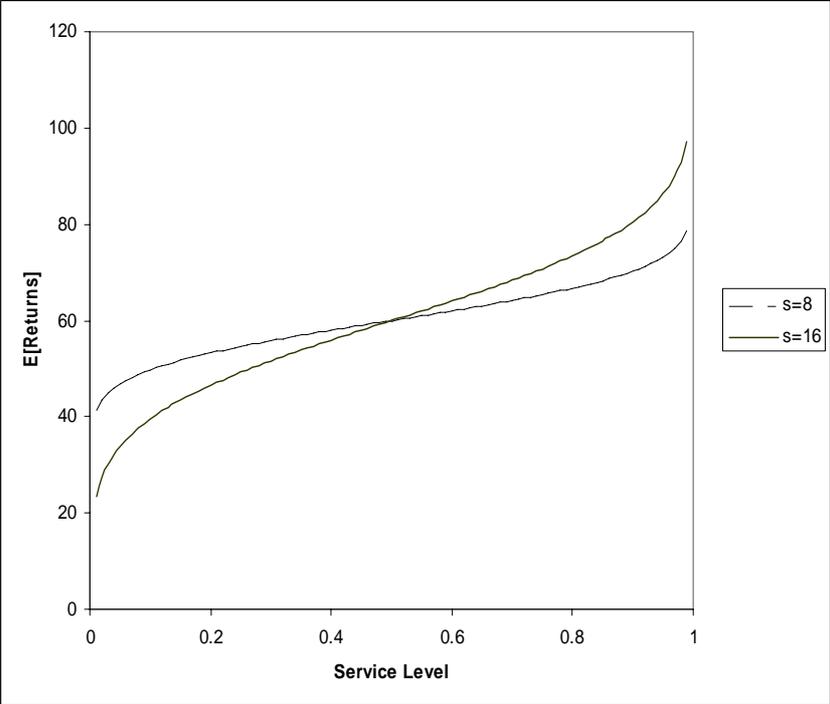


Figure 2 Heteroskedasticity Significance of White Test Dummy Coefficients

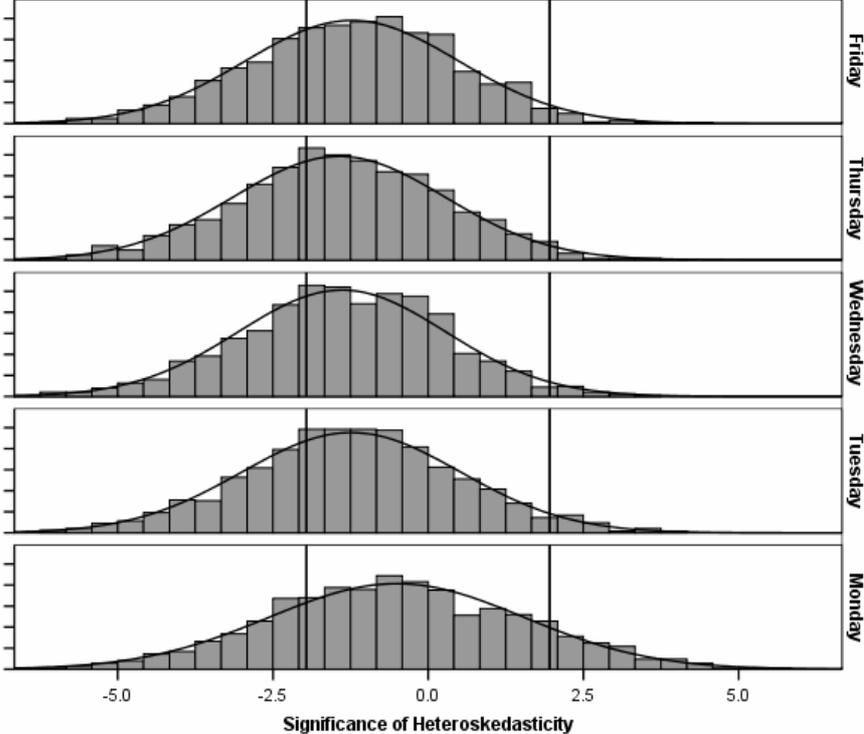


Figure 3 Parameter Significance for Equation (6)

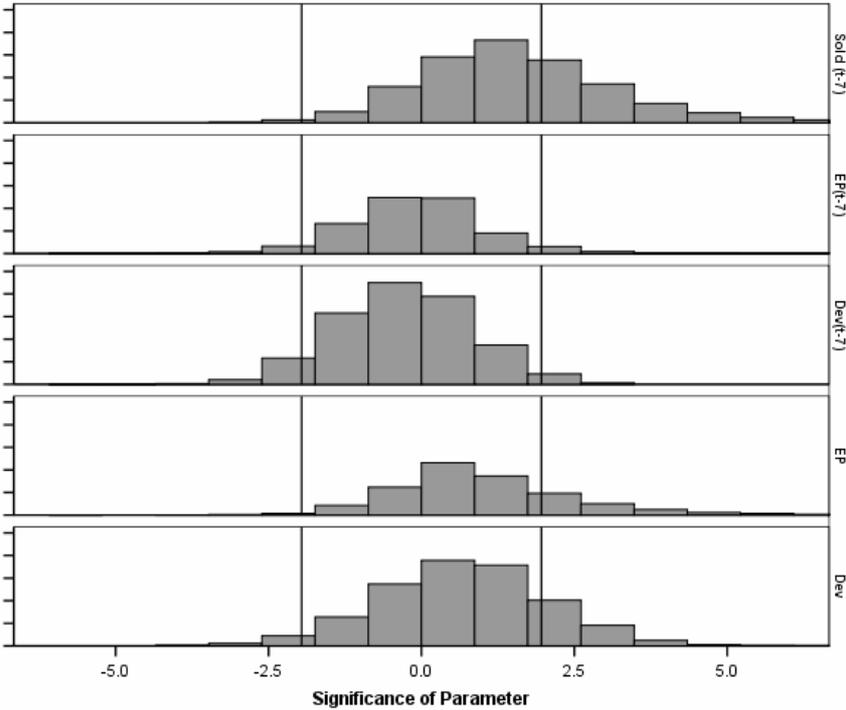


Figure 4 Coefficients of Determination

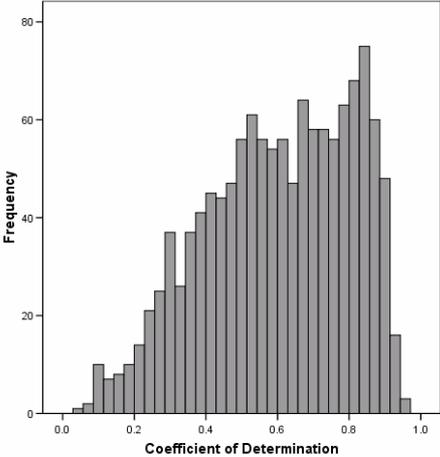


Figure 5 Normal Probability Plot of Residuals

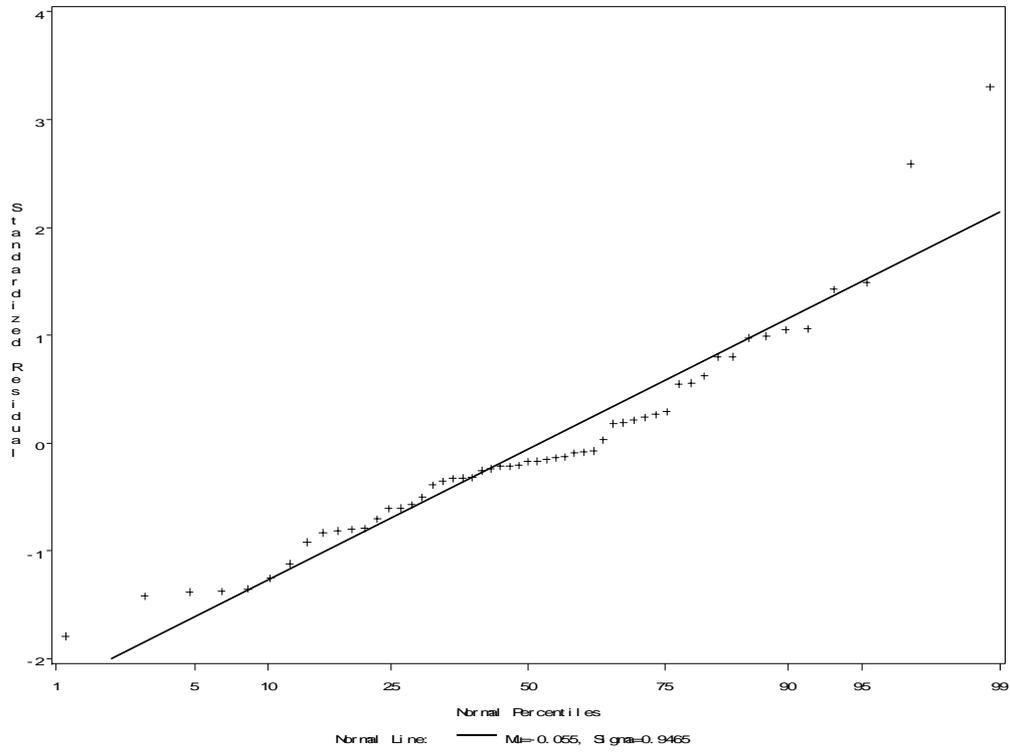


Figure 6 Kolmogorov-Smirnov P-value versus Average Sales

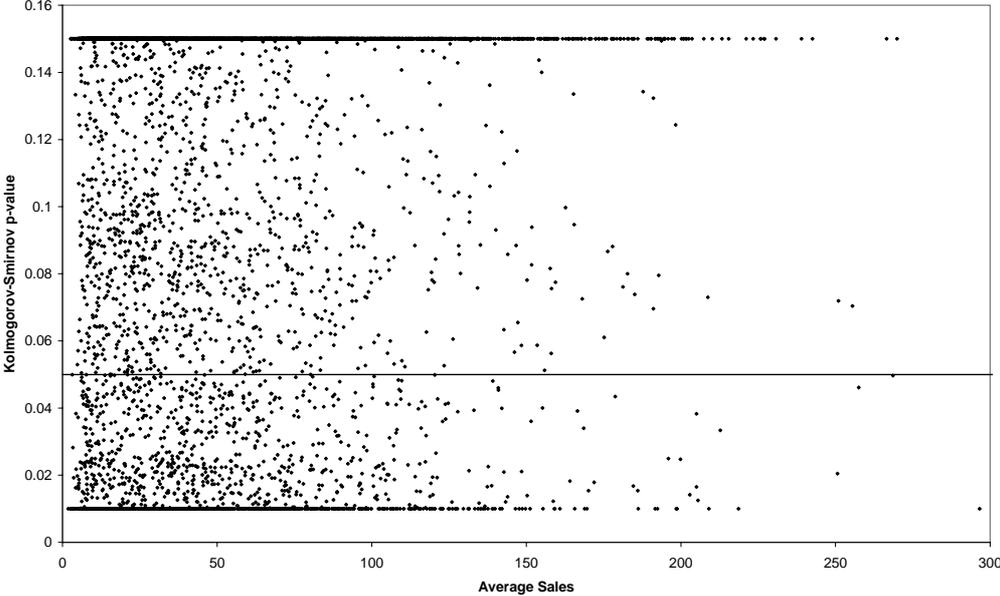


Table 1 Significance of Parameters in Equation (4)

Parameter	Frequency of Significance	Frequency of Estimates	Significance Percentage
Intercept	1273	1274	100%
Monday	1139	1257	91%
Tuesday	1185	1262	94%
Wednesday	1186	1273	93%
Thursday	1190	1267	94%
Friday	1202	1269	95%
$D^P_{id,i}$	666	1274	52%
$S^c_{id,i-6}$	1018	1274	80%
$D^P_{id,i-6}$	282	1274	22%
$D^R_{id,i-6}$	120	1152	10%
$D^R_{id,i}$	572	1151	50%

Table 2 Significance of Parameters in Equation (6)

Parameter	Frequency of Significance	Frequency of Estimates	Significance Percentage
Intercept	7572	7602	100%
$D^P_{id,d,i}$	2055	7600	27%
$S^c_{id,d,i-6}$	3424	7602	45%
$D^P_{id,d,i-6}$	1086	7602	14%
$D^R_{id,d,i-6}$	484	4012	12%
$D^R_{id,d,i}$	1125	3918	29%

Table 3 Expected Deliveries, Shortages and Returns

Service Level		Equation (4)		Equation (6)	
		Backtracking	Theoretical	Backtracking	Theoretical
97%	Deliveries/SKU/day	64.99	64.98	64.07	64.05
	Deliveries/day	82796	82781	81623	81602
	Shortage/SKU/day	0.15	0.087	0.12	0.081
	Shortage/day	192	111	154	104
	Returns/SKU/day	14.58	14.20	13.63	13.26
	Returns/day	18579	18092	17368	16894
93%	Deliveries/SKU/day	61.95	61.94	61.23	61.21
	Deliveries/day	78923	78909	78008	77987
	Shortage/SKU/day	0.26	0.23	0.22	0.22
	Shortage/day	326	296	284	276
	Returns/SKU/day	11.65	11.31	10.90	10.56
	Returns/day	14839	14405	13883	13451

Table 4 Delivery Policy for a 97% Service Strategy

Days	Observed	Equation (6)	Difference	Number of Days
Normal Days	63.6	64.3	0.8	453288
Days with D ^P	71.9	68.0	-3.9	115283
Days with D ^R	88.5	88.3	-0.3	7733

Figure Captions

Figure 1 Expected Returns as a Function of Service Level

Figure 2 Heteroskedasticity Significance of White Test Dummy Coefficients

Figure 3 Parameter Significance for Equation (6)

Figure 4 Coefficients of Determination

Figure 5 Normal Probability Plot of Residuals

Figure 6 Kolmogorov-Smirnov P-value versus Average Sales

Table Headings

Table 1 Significance of Parameters in Equation (4)

Table 2 Significance of Parameters in Equation (6)

Table 3 Expected Deliveries, Shortages and Returns

Table 4 Delivery Policy for a 97% Service Strategy