The length of stay determinants for sun-and-sand tourism: An application for the Region of Murcia

Artal Tur, Andrés Andres.artal@upct.es
García Sánchez, Antonio A.garciasanchez@upct.es
Departamento de Economía
Sánchez García, Juan Francisco jf.sanchez@upct.es
Departamento de Métodos Cuantitativos e Informáticos
Universidad Politécnica de Cartagena

RESUMEN

While tourist arrivals increase annually in Spain, tourist average real expenditure has decreased significantly over the last few years, with important effects on tourism revenues. The process is clearly driven by the reduction of the length of stay of tourists at destinations, but surprisingly this variable has received little attention in the literature. We estimate a length of stay function for sun-and-sand tourists visiting the Region of Murcia over the period 2002-2006 using count data models. Our results show that both tourists' personal and family characteristics together with economic variables (budget restrictions, income and prices) are key factors in determining the duration of the stay. Quantitative identification of the determinants of a tourist's length of stay could provide important guidelines for designing policies aimed at influencing length of stay in tourist's seaside destinations.

***Palabras clave*:**

Length of stay, count data models, socio-economic profiles, holiday's choice.

***Clasificación JEL (Journal Economic Literature)*:**

C12, R11, R58

***Área temática*:** 8. Turismo y Métodos Cuantitativos

1. INTRODUCTION

Spain is the number two destination in the world in terms of tourist arrivals and tourism revenues generated by inbound tourism, receiving around 60 million visitors (just behind France with 79 million) and 51 million dollars in income in 2007 (just behind the US with 85 million), respectively. However, the tourism sector is now facing important changes. While tourist arrivals have increased 3.8 % annually over the period 2000-2007, total real tourist revenue have remained relatively constant, which has led to a decrease of 22% in tourist average real expenditure since the beginning of this century (Tourism Studies Institute (IET)). The reduction of tourist average expenditure is a two-fold process, with daily real expenditure per tourist increasing slightly over the years, but the length of stay falling appreciably. According to the *Hotel Occupancy Survey* of the National Statistics Institute (INE) for Spain, there has been a cumulative decrease of 16% in the average stay of a tourist since 2000, which means an annual reduction of 2.4%. This process is not unique to Spain, but constitutes an important characteristic shared by the world's foremost tourist destinations today. In this sense, length of stay emerges as one of the key variables when explaining the trend in tourist expenditure worldwide (World Tourism Organisation, 2007). Surprisingly, studies on the determinants of the length of stay have not prolifered in the literature. In this context, our research abounds in the study of the determinants of length of stay, estimating a function of the determinants of length of stay for sun-and-sand tourists with data from the Spanish Region of Murcia from 2002 to 2006.

The remainder of the paper is organized as follows. In Section 2, a review of the previous literature on studies of length of stay determinants is carried out. Section 3 includes a description of the main characteristics of the area under study. In Section 4, the economic model is presented, while Section 5 describes the econometric aspects of the empirical model. Section 6 outlines the results of the estimation and Section 7 summarizes the conclusions of the research.

2. LITERATURE REVIEW

As stated above, the literature on tourist' length of stay determinants is still incipient. Regarding the few specific studies existing on this issue, we must note that most of them proceed in a descriptive way, just analysing the relationship between the duration

of the stay and the socio-economic profile of the tourist (Sung et al., 2001; Seaton & Palmer, 1997). Other contributions enrich the study through adding new explanatory variables, such as the family life cycle position or time transport distance between the country of origin and destination (Bernham & Kim, 1999; Oppermann, 1995). Subsequently, we find a small group of papers that apply econometric techniques as the pioneer contribution by Mak, Monkur and Yonamine (1977) for US tourist in the Hawaii Islands.

More recent contributions include Bell and Leeworthy (1990), who develop an exercise for modelling the consumer demand of Florida Beach visitors, obtaining the expected signs for daily-cost-at-destination (negative coefficient) and tourist income (positive coefficient), and Fleischer and Pizam (2002), who apply a tobit model to analyse the length of stay of a cross-section of Israeli tourists, obtaining that income and some individual characteristics (age and health status) are positively related to the duration of their stay. Finally, Alegre and Pou (2006) estimate a length-of-stay function through a binomial logit for the Balearic Islands, identifying the influence of individual and economic constraints in determining the duration of the stay.

3. DESCRIPTIVE EVIDENCE

Our database comes from a survey on sun-and-sand tourists conducted by the Department of Tourism, Commerce and Consumption of the Murcian Regional Government, which compiles information on tourist expenditure, accommodation prices and socio-economic profiles of visitors, including their length of stay, type of accomodation, country of residence and season of the visit, among others. The length of stay in the Region of Murcia (hereafter referred to as Murcia) was 11.6 days in 2006 according to our survey data. The average stay of tourists arriving to the region declined by 22% in absolute terms over the period 2000-2007 (19% for Spanish tourists and 34% for foreign tourists), which reports an annual average decrease of 3.4%

A first look at our database reveals that tourists arriving to Murcia's coast predominantly choose a rent apartment or villa (58%) as their type of accommodation. The tourists who select a hotel as accommodation account for the remaining 42% of the survey, with 33% preferring high-class hotels (5, 4 or 3-star) and 9% lower-class hotels. In terms of their countries of residence, these sun-and-sand tourists come mainly from

the rest of Spain (68.5%), particularly from Andalusia, Valencia and Madrid. Visitors from the rest of the world, European Union citizens basically, add another 26% to this figure and residents in Murcia itself account for the rest (around 5.5% of the sample). Tourists come with their families in the majority of cases (73%), while 23% come with friends and they normally arrive to this coast in high-season months (71% of all cases). First-time visitors represent just 36% of the total sample, while tourists that make more than one trip a year account for 44% of the whole sample.

4. ECONOMIC MODEL

The study relies theoretically on a *consumer behavior model* approach, with a typical consumer that maximizes a utility function, including q, the vector of consumer goods (excluding tourist services), z, the vector of characteristics that define the holiday choice, and t, the length of the holiday (Dubin & McFadden, 1984; Hanemann, 1984). This length can be separated into two components, the first one is the travel time to destination, t_{trans}, and the second is the length of stay at destination proper, t_{tour}, so total length is defined as $t = t_{trans} + t_{tour}$. J represents alternative holidays. The utility function also includes a vector of taste shifters, η, with characteristics of the consumer (tourist) that might influence his or her preferences and a random term, ε, for non-observable characteristics of the trip (McFadden, 1981). The consumer determines the values of q, z and t that maximize his or her utility function:

$$\max_{q,\, z\,,\, t_{trans},\, t_{tour}} \; U\,(q,\, z,\, t_{trans},\, t_{tour},\, \eta,\, \varepsilon\,) \qquad (1)$$

subject to income restrictions, Y, and time constraints, T:

$$pq + p_{trans} + p_{tour}\, t_{tour} \leq Y,$$

$$t_{trans} + t_{tour} \leq T,$$

$$q,\, z,\, t_{trans},\, t_{tour},,\, p,\, p_{trans},\, p_{tour} \geq 0$$

where p is the vector of prices of q, and p_{trans} and p_{tour} are the travel price and holiday price per day at the destination, respectively. The latter is determined by the characteristics of the holiday.

Following Pollak (1969), the conditional demand function for the length of stay at destination j can be expressed as:

$$t_{j\text{-}tour} = t_{tour}\,(p,\, p_{j\text{-}tour},\, z_j,\, Y - p_{j\text{-}trans},\, T - t_{j\text{-}trans},\, \eta,\, \varepsilon_j) \qquad (2)$$

Expression (2) conditions the length of stay on the selected holiday characteristics and income and time constraints. What makes this conditional demand function appealing is the fact that the demand function for length of stay can be estimated taking pre-assigned values for the choice of destination and the set of holiday characteristics. Equally, we assume weak separability between the tourist trip and consumer goods other than tourism, q. Weak separability implies that income and the prices of goods outside the tourist trip only affect the demand function through their effect on total expenditure on the trip. Under this assumption, the conditional demand function for length of stay can be written as:

$$t_{j\text{-}tour} = t_{tour}\,(\,p_{j\text{-}tour},\, z_j,\, Y - pq - p_{j\text{-}trans},\, T - t_{j\text{-}trans},\, \eta,\, \varepsilon_j) \qquad (3)$$

Thus, in our theoretical model the length of stay is a function of daily price of the holiday (dependent upon its selected characteristics), holiday characteristics, total expenditure available for the holiday, time available for it, characteristics of the tourist and a non-observable random effect.

5. ECONOMETRIC ISSUES

Figure 1 plots the distribution of the total length of stay, with a sample mean value of 11.6 days, a maximum value of 31 days and a minimum of 1[1]. Tourists staying for exactly 1 or 2 weeks represent around 23% of total tourists, with data showing four attractor points on values of 7, 10, 15 and 30 days, and an important presence of tourists located between 1 and 15 days of stay.

[Insert Figure 1 here]

Our dependent variable is composed by nonnegative integer values, so we decide to apply count data models to study its main determinants, given that this kind of models are very well suited to deal with this characteristics of the dependent variable. Applying this method of estimation will allow us to observe the estimated values of elasticities at a given point of the sample and not just for the intervals of the dependent variable, as in discrete choice models, subsequently improving the information obtained by the estimation process and thus enriching the discussion of associated results (Wooldridge, 2002).

[1] Our dependent variable then ranges from 1 to 31 days, an interval that we deem the maximum time available for holidays as a working assumption in accordance with the theoretical model.

The database contains information for each tourist on: 1) length of stay (LOS) (in days), 2) age: groups of under 25 years, 26-45, 46-55, 56-65, over 65 years, 3) nationality: resident in Murcia itself, resident in the rest-of-Spain, resident in the rest-of-the-World, 4) type of accommodation: rent apartment & tourist villa, high-class hotel (5,4 and 3-star), low-class hotel (other hotels), 5) company for the holidays: alone, with friends or with the family, 6) whether this is the tourist's first visit to this destination or not, 7) whether or not the tourist makes just one trip per year, 8) tourist expenditure (including expenditures in the country of origin and transport costs) and 9) daily price of accommodation. Variables 8) and 9) were specified as continuous variables. We only have information on accommodation prices in our database, so this is going to be the price variable of our model.[2] The rest of the variables are specified as dummies, taking one of the categories as a reference. So, the specification of the length of stay function is as follows:

$$LOS_i = \beta_1 + \beta_2 AGE_i + \beta_3 NATION_i + \beta_4 ACCOMOD_i + \beta_5 COMPANY_i +$$

$$+ \beta_6 FIRSTVISIT_i + \beta_7 VARIOUSTRIPS_i + \beta_8 TOUREXP_i + \beta_9 DPA_i + \varepsilon_i$$

$$i = 1\text{-}1915.$$

6. RESULTS

We use the STATA 10.0 software to conduct our estimation process and start estimating the model as if the dependent variable followed a Poisson distribution. Nevertheless, the mean of the dependent variable is 11.6 with variance of 52.1, which clearly signals the presence of overdispersion in the distribution, with $E(Y) \neq V(Y)$. Pearson and χ^2 tests also reject the null hypothesis of the dependent variable following a Poisson distribution.[3] We then opt for using a negative binomial as the underlying distribution, estimating the model by QML with robust standard errors and covariates,

[2] To this respect, we are aware that this variable is not strictly the total daily price of the holidays, but rather just a part of it, but it is the only proxy that we have available in our dataset. One advantage of using this price variable is that accommodation price is, rigorously speaking, an exogenous variable in our specification, while other authors have proxied the daily price of holidays through the use of the daily expenditure of the tourists, which is not a price variable and also brings important endogeneity problems into the estimations. As a result, we follow the theoretical approach of the model more closely.

[3] We do not show these results for space constrains, but they are available by request to the authors.

which ensures consistency in the coefficients of the model (Wooldridge, 2002; Cameron & Trivedi, 1998).

Estimation results for negative binomial are shown in Table 1A, with four columns containing calculated elasticities, standard errors and statistical significance (z-statistics and p-values) of the explanatory variables in the model. In the case of dummies, estimated elasticities are evaluated at the value 1 of the variable and for continuous variables we evaluate them at their mean values (Table 1A), as well as at other values interesting for the research (Table 1B). Information criteria and other measures of fit are also at the bottom of Table 1A (Log-L, AIC, Schwartz criteria, McFadden Adj. R^2). The category of reference in the general model is one individual (tourist) aged over 65, resident in Murcia, accommodated in a high-class hotel, coming with his or her family not for the first time to this destination and making just one trip per year. In the following paragraphs we include detailed results for each explanatory variable in the general model.

Age: Estimations show that being a tourist aged 65 or less reduces the length of stay with respect to the over 65 group. However, the coefficients of the age groups between 46-55 and 56-65 years are not statistically significant, reflecting perhaps that these three groups of older tourists normally spend their holidays together as a family. Visitors aged between 26-45 years record an elasticity value of -11% with respect to the reference category and those aged 25 years or less, a value of -10.9%, indicating that in general terms, younger tourists stay for around 10% less. These results are in line with those of Seaton and Palmer (1997) and Alegre and Pou (2006), although we observe a higher incidence in reduction of average stay for younger visitors and a lower incidence for older tourists in comparison with weekly-package holidays in the Balearic Islands.

Nationality: The category of reference are now those tourists who live in Murcia itself, so our results show that being a resident in the rest-of-Spain reduces the stay by 29% with respect to the reference category, meanwhile if a tourist comes from the rest-of-the-World, obtained elasticity is -24%. In both cases, this is an important result for this type of destination, showing that distance significantly influences the duration of the stay.

Accommodation: This is another important issue in this research, given that one of our aims was to focus on the differences in stay determinants between destinations where a large number of tourists come to hotels versus destinations where an important

complex of apartments and villas for rent complement this accommodation supply. In our case, the results show that tourists choosing a rent apartment or a villa stay some 30% longer than those staying in a high-class hotel and those in a low-class hotel stay 5% less time than those in the reference category. Differences seem to be significant and will be discussed in more detail in the following sections. Alegre and Pou (2006) obtain the same sign for this type of accommodation in their study, although we cannot compare elasticities because they use a discrete choice econometric approach.

Company: Tourists who come with their family are the reference category, so we can observe that belonging to the other two categories of company reduces the length of stay: coming with friends does so by 12.8% and coming alone by 6% (although this coefficient does not appear to be statistically significant).

First Visit: We can see that sun-and-sand tourists coming for the first time to this coast stay for a shorter time (7.6% less) than those with a greater knowledge of the destination.

More than one trip: tourists that make more than one trip per year also register shorter lengths of stay (-8.7%) in comparison with those who travel just once a year.

Tourist Expenditure: For general tourists, the elasticity associated to the income variable stands at 19%, which seems to be a small value, considering that we used to think of tourism as an above average income earning or luxury good.[4]

Daily price of accommodation: Our results show that the price variable has an important effect on holiday duration, with an estimated elasticity of -61%, emphasizing as in other studies the importance that prices have on tourists' travel decisions. The estimated value of price elasticity is lower than in Alegre and Pou (2006), who obtain a value of -1.79, although differences may arise because they use the total daily cost of the holiday and we have to use the price-per-day of accommodation. In contrast, our results are far higher than those obtained by Bell and Leeworthy (1990), with an estimated price elasticity of -0.15, and Mak et al. (1977, 1979) with an elasticity of around -0.25.

[4] For example, Alegre and Pou (2006) found an estimated elasticity of 2.39 for the total holiday expenditure variable, but they restrict their sample to German and British tourists visiting the Balearic Islands. Our results are more in line with those of Bell and Leeworthy (1990) for Florida Beach tourists, who obtain an income elasticity of 0.37 in their exercise with micro data.

In light of the relevance that the last two variables (income and prices) have in studies of tourist demand, we decided to evaluate them at different levels of the explanatory variable, keeping the rest of the variables constant (*ceteris paribus*). Table 2B allows us to observe that tourists with lower income, proxied by expenditure, also record lower income elasticities, while tourists with higher income display correspondent greater income elasticity.[5] Equally, estimation results show that tourists paying a higher price for accommodation per day are clearly more sensitive to variations in price levels when they have to choose the length of stay of their holidays. Nevertheless, results for high income tourists, with an income elasticity of 0.33 and a price elasticity of -1.93, seem to be closer to those obtained by Alegre and Pou (2006) for German and British tourists visiting the Balearic Islands.

7. CONCLUSIONS

Worldwide tourism demand is evolving towards an increase in the number of annual trips, together with an important reduction in the duration of the stays. Short trips around the world represent an important part of total demand, while traditional sun-and-sand destinations have to deal with the decrease in the average stay of their visitors. Real tourist average expenditure has been declining in Spain since the beginning of the century, driven precisely by this significant reduction in the average length of stay of tourists. In this context, new strategies are necessary in order to improve or even maintain the revenues of tourist countries and, consequently, more information on length of stay determinants is needed in order to guide such policies.

In this context, our research has been directed to identify key variables determining the length of stay of tourists from a microeconomic point of view, as a way of extending this still incipient literature. Pursuing this objective, a length of stay function has been estimated for sun-and-sand tourists in the Spanish Region of Murcia for the period 2002-2006. Results had show how relevant tourists' personal and family characteristics together with tourist's economic constrains, such as level of income and the price of goods and services consumed, determine the duration of their stay, as the

[5] The intuition behind this result seems to be counterintuitive, because we would expect higher income-elasticity for lower-income tourists. However, this result would be expected if the basket of consumed goods and services were the same for both groups of tourists, which does not seem to be the case, as we will discuss further. In this case, the rationale behind this result surely has to do with the kind of "holiday choice" that high-income tourists make in comparison to low-income visitors.

theoretical model pointed out. More specifically, variables such as nationality or country-of-origin of tourists, age, company for holidays and especially accommodation price have been identified as the main factors influencing the length of a tourist's stay.

The new econometric approach employed in estimating microeconomic determinants of a tourist stay, through the use of count data models, has proved more flexible and informative in this type of exercise in comparison with discrete-choice methodology, especially in the case of binomial logit models. Following this approach, we have been able to identify the presence of significant differences in price and income elasticities for high-income and low-income tourists as another important result of the investigation. In summary, our results have been able to show, and what is more important, to quantify, the fact that different holiday choices linked to different groups of tourists induce important differences in the determinants of the length of stay. This general result clearly advocates for the need of bearing these differences in mind when designing policies directed to extend the tourists' length of stay at destinations.

8.- REFERENCES

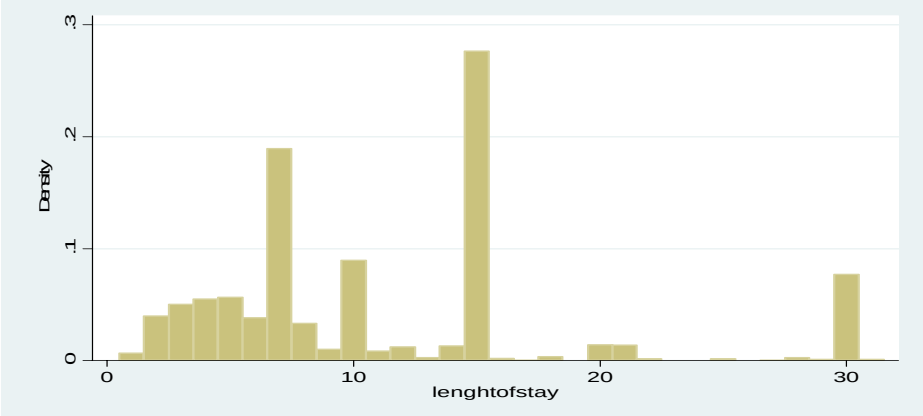
- ALEGRE, J., and POU, L. (2006). "The length of stay in the demand for tourism". Tourism Management, 27, pp. 1343-1355.
- BELL, F. W., and LEEWORTHY, V. R. (1990). "Recreational demand by tourists for Saltwater Beach Days". Journal of Environmental and Management, 18, 3, pp.189–205.
- BERMAN, M. D. and KIM, H-J. (1999). "Endogenous on-site time in the recreation demand model". Land Economics, 75, 4, pp.603–619.
- CAMERON, A.C., and TRIVEDI, P.K. (1998). Regression Analysis of Count Data. Cambridge (Mass.): Cambridge University Press.
- DEATON, A. S., and MUELLBAUER, J. (1980). Economics and consumer behaviour. Cambridge: Cambridge University Press.
- DUBIN, J. A., and MCFADDEN, D. L. (1984). "An econometric analysis of residential electric appliance holdings and consumption". Econometrica, 52, 2, pp.345–362.
- FLEISCHER, A., and PIZAM, A. (2002). "Tourism constraints among Israeli Seniors". Annals of Tourism Research, 29, 1, pp.106–123.

- HANEMANN, W. M. (1984). "Discrete/continuous models of consumer demand". Econometrica, 52, 3, pp.541–561.
- HECKMAN, J. J. (2001). "Micro data, heterogeneity, and the evaluation of public policy: Nobel lecture". Journal of Political Economy, 109, 4, pp.673–748.
- LI, G., SONG, H., and WITT, S. F. (2006). "Time varying parameter and fixed parameter linear AIDS: An application to tourism demand forecasting". International Journal of Forecasting, 22, pp.57–71.
- LIM, C. (1997). "An econometric classification and review of international tourism demand models". Tourism Economics, 3, pp.69–81.
- MANGION, M. L., DURBARRY, R., and SINCLAIR, M. T. (2005). "Tourism competitiveness: Price and quality Tourism competitiveness: Price and quality". Tourism Economics, 11, pp.45–68.
- MAK, J., MONCUR, J., and YONAMINE, D. (1977). "Determinants of visitor expenditures and visitor length of stay: A cross-section analysis of US visitors to Hawaii". Journal of Travel Research, 15, pp.5–8.
- MCFADDEN, D. (1981). "Econometric models of probabilistic choice". In Manski,. and McFadden, D. (Eds.) Structural analysis of discrete data, pp. 198–272. Cambridge (Mass): MIT Press.
- OPPERMANN, M. (1995). "Travel life cycle". Annals of Tourism Research, 22, 3, pp. 535–552.
- POLLAK, R. A. (1969). "Conditional demand functions and consumption theory". The Quarterly Journal of Economics, 83, 1, pp. 60–78.
- SEATON, A. V., and PALMER, C. (1997). "Understanding VFR tourism behaviour: The first five years of the United Kingdom tourism survey". Tourism Management, 1, 6, pp. 345–355.
- SUNG, H. H., MORRISON, A. M., HONG, G.-S., and O'LEARY, J. T. (2001). "The effects of household and trip characteristics on trip types: A consumer behavioural approach for segmenting the US domestic leisure travel market". Journal of Hospitality & Tourism Research, 25, 1, pp. 46–68.
- WOOLDRIDGE, J. (2002). Econometric Analysis of Cross Section and Panel Data. Cambridge (Mass): MIT Press.

- WORLD TOURISM ORGANISATION (2007). World Overview & Tourism Topics. Madrid: United Nations World Tourism Organisation (UNWTO).

TABLES AND FIGURES

Figure 1. Length of stay at the coast of Murcia 2002-2006



Source: Own elaboration on Regional Government of Murcia survey data.

Table1. A) Elasticities after negative binomial regression

variable	ey/ex	Std. Err.	z	P >|z|	X value
various trips	-0.0870	0.015	-5.69	0.000	1
rest-of-Spain	-0.2916	0.027	-10.63	0.000	1
rest-of-the-World	-0.2433	0.029	-8.25	0.000	1
until25 years	-0.1091	0.034	-3.18	0.001	1
bt2645 years	-0.1100	0.027	-4.08	0.000	1
bt4655 years	-0.0352	0.030	-1.15	0.251	1
bt5665 years	-0.0486	0.032	-1.51	0.130	1
alone	-0.0682	0.042	-1.59	0.112	1
friends	-0.1282	0.018	-7.04	0.000	1
first visit	-0.0763	0.016	-4.75	0.000	1
low-class hotel	-0.0507	0.025	-1.98	0.047	1
rent aprt	0.3000	0.017	16.84	0.000	1
dpa	-0.6122	0.018	-32.85	0.000	31.6
totexp	0.1926	0.020	9.46	0.000	567.2
Log-L	**AIC**	**Schwartz**	**Mc-Fadden Adj-R^2**		
-4 907.39	9 846.78	9 935.70	0.208		

B) Elasticities after negative binomial regression: selected values of explanatory variables (dpa, totexp)

variable	ey/ex	Std. Err.	z	P >|z|	X value
dpa1	-0.6122	0.018	-32.85	0.000	31.6
totexp1	0.0849	0.008	9.46	0.000	250
dpa2	-0.6122	0.018	-32.85	0.000	31.6
totexp2	0.3396	0.035	9.46	0.000	1000
dpa3	-0.2420	0.007	-32.85	0.000	12.5
totexp3	0.1926	0.020	9.46	0.000	567.2
dpa4	-0.4841	0.014	-32.85	0.000	25
totexp4	0.1926	0.020	9.46	0.000	567.2
dpa5	-0.9683	0.029	-32.85	0.000	50
totexp5	0.1926	0.020	9.46	0.000	567.2
dpa6	-1.936	0.058	-32.85	0.000	100
totexp6	0.1926	0.020	9.46	0.000	567.2