Determinants of leaving home during the pandemic: can the mood be a hidden driver?

Anar Aghazada, Ekaterina Petrova, Mariona Sales

June 9, 2022

Abstract

In this study, we use a survey data from Catalonia during the Covid-19 lockdown, and analyse the effects of different factors on people's willingness to go out. Beyond the conventional paradigm, the paper specifically focuses on the effects of the mood, worries, and thoughts on this willingness.

1 Objective

On the 14^{th} of March 2020, the Spanish government decreed a state of alarm throughout Spain and a complete lockdown for the whole population started in order to limit the spread of the COVID-19 pandemic. All activities related to social contact were suspended, including the closure of much of the industrial, commercial, cultural, restaurant, leisure, and sportive activities. Staying at home was the most appropriate and responsible behaviour individuals could have to stop the transmission of the virus and protect the health of the most vulnerable people.

Staying at home for a long period of time could alter life habits and create psychological discomforts. In this study, we analyse how depression and anxiety affected the number of days Catalan people were outside the house during March and April 2020, the first period in which society started to live with COVID-19 pandemic situation.

We also want to examine how worries about health and the economy affected the number of days people went outside. In terms of health, it was understandable that people were worried about their own health but also society's health due to the increase in infections, admissions to Intensive Care Units at hospitals and the number of deaths that were detected in that period. An economic crisis was starting and supply chains were disrupted so people were worried about the consequences it could have to their own economy.

We should need to consider that citizens were able to go out for priority reasons such as shopping, walking the dog or taking care of dependent people, but these activities were also perceived as alleviating the fact of being at home. One of our objectives is also to study if people took advantage of it in an irresponsible way.

In the next sections, we are going to talk about the theoretical framework, the econometric specification, main issues related to the data, main results and conclusions, weak aspects of the paper and improvements.

2 Literature review

We could find related literature only focused on studying the level of depression and anxiety and predictor of it. In the study "Longitudinal evaluation of the psychological impact of the COVID-19 crisis in Spain" Planchuelo-Gómez et al. (2020) authors compare the level of depression, anxiety, stress and other in two periods – March 28-April 5 and April 28-May 15. It turns out that age, sex, if person health worker or not, employment activity and marital status have significant effect on depression level. Also the article provides the graph (Fig. 1) with restrictions on activity in Spain, daily cases of Covid and deaths in our period. We can see that our period of survey is homogeneous on restrictions and death cases but has a peak of infection cases.

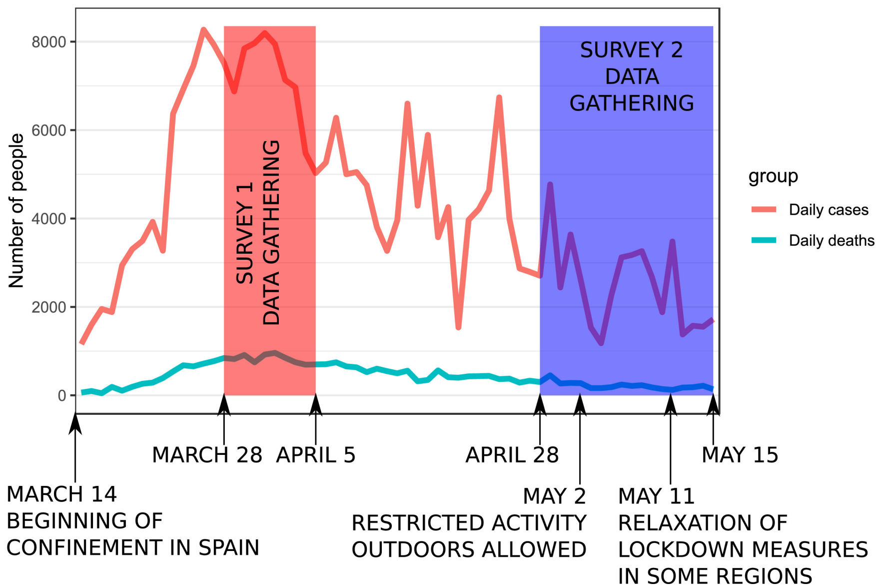

Figure 1: Evolution of daily COVID-19 confirmed positive cases and deaths in Spain from the beginning of the confinement to the final date of data gathering. Source: Planchuelo-Gómez et al. (2020)

The other work "The impact of the COVID-19 lockdown on depression sufferers: a qualitative study from the province of Zaragoza, Spain" Aguilar-Latorre et al. (2022) made some qualitative conclusions about the effect of leaving the house on depression level: "The physical space (availability and quality) of the home was an important factor shaping well-being".

"Coping behaviours associated with decreased anxiety and depressive symptoms during the COVID-19 pandemic and lockdown" Fullana et al. (2020) is the other Spanish study that brought out that being outdoors or looking outside were the best predictors of lower levels of depressive symptoms.

From the literature review we can conclude that there could be simultaneity bias in our model because there is evidence that inability to go outside affects the level of depression. But as long as we analyse only one period that has the same restrictions among all the timelines, there are reasons to think that depression at the same time has an effect on leaving the house, we will just try to control it and make a model.

3 Data

The data set that has been used comes from different surveys done by CERES to the Catalan population between the 21^{st} of March and the 3^{rd} of May of 2020 concerning how they were living the lockdown, the state of alarm and the crisis of COVID-19.

The data set contains 4,574 observations which were interviewed in one of the 4 periods[1] that the study establishes. These dates do not coincide with the waves that the Catalan government has set up during these two years of the pandemic to impose the corresponding restrictions.

In the distribution of the sample, we can find an overrepresentation of individuals living in Tarragona province (58,2%). Concretely, the municipalities of Reus, Tarragona, Amposta and Torrefarrera actively participated in the survey. In order to tackle this issue and obtain a representative sample, the weights for the variables sex, age and province have been restored according to the distribution of the population in Catalonia based on IDESCAT data. After dropping out the observations which have NA values in any of the columns, we are left with a sample of 1,945 observations which turns out to only be part of waves 2 and 4.

The data set has 57 columns. Our dependent variable is the number of days individuals have gone out of home in a week during the COVID-19 lockdown. In terms of the explanatory variables, the one that we have used are those related to sociodemographic characteristics: age, sex, province, labor situation and whether there are children or no in the house. There are also variables related to the reasons why the interviewees have left home as well as ordered variables collecting the different moods individuals have faced during lockdown from which summary variables for sadness-depression, anxiety, anger-hostility and happiness have been created. We have also used some variables for worriedness about their own and society health as well as the status of their own economy. Finally, there is a numeric variable related to the people's perception of days that the lockdown will last. For each categorical level, we have created dummy variables.

[1]Dates of the waves: 1^{st} from 21^{st} of March to 24^{th} of March 2020, 2^{nd} from 27^{th} of March to 31^{st} of March 2020, 3^{rd} from 21^{st} of March to 24^{st} of March 2020, 4^{th} from 3^{rd} of April to 7^{th} of April 2020 and 5^{th} from 27^{th} of April to 14^{th} of April 2020.

Descriptive statistics					
Variable	Obs	Mean	Std. dev.	Min	Max
Days	1,945	2.4	2.2	0	7
Age	1,934	47	12	14	84
Sadness&depression	1,945	4.13	2.4	0	10
Anxiety	1,945	4.4	2.5	0	10
Perception days	1,933	33.3	15.2	0	201

Our dependent variable is a count variable that has the following characteristics:

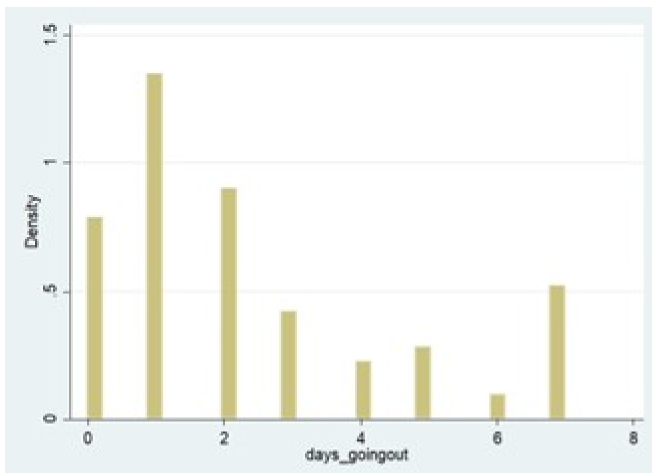

Figure 2: Density function of number of days going out

- Non-negative: it is a limited dependent variable since it cannot take negative values.
- Significant proportion of zeros: there is 17.17% of our sample for which the observed value of days going out is zero, which represents a significant proportion of the observations.
- A reduced number of integer values: the number of values that this variable can take is low because the period of time studied is a pandemic in which going out was limited. As soon as we mode up, the relative frequency of the number of observations is reduced.

The Poisson or Negative Binomial distribution is the adequate one for these specific features that our dependent variable has.

4 Econometric specification

Here, our main target variable is the count of days per week that people went out. We model this variable in two different ways, first as a binary variable, which takes 0 if the individual has not gone out at all, and takes 1 if he/she has. Then, we also have the exact count of days, to see the effect of the corresponding explanatory variables on this.

As it is seen, we need two types of models to construct our study. For the first variable, we will need to choose a binary choice specification to model the effects of the regressors on the probability of going out. As the linear probability model is deficient for having lower than 0 or higher than 1 values, our main options include Logit and Probit models. For the count data, we can use either Poisson or Negative Binomial, depending on some features of the data that we will talk about later. In literature, that kind of study is mainly done with either double-hurdle model or zero-inflation models. The first part of the double-hurdle is logit and the second part is truncated Poisson or negative binomial. Note that, truncated models use only the individuals whose target value is higher than 0. On the other hand, if the data has an excess amount of zero, or if the possible reasons for zero are more than one, zero-inflated models are preferred. As for not having an excessive number of zero, we first try to build a double-hurdle. However, the reasons for zero can be various, whereas we model a Covid data some people personally can choose to stay at home without a government restriction. Thus, we will try to capture a possible effect with a zero-inflated model afterward.

For the first part of the double-hurdle, our choice is using Logit, which made a slightly better job than Probit when was compared using log-likelihood values. Below, you can see the mathematical representation.

$$D_i = F(X_i^{'}\alpha_1 + Sad\&Dep_i\alpha_2 + Anxiety_i\alpha_3 + W\_ownh_i\alpha_4 + W\_societyh_i\alpha_5 + Perception_i\alpha_6 + \epsilon_i)$$

Where, F is a probability function, $X^{'}$ term shows the individual characteristics, like gender, age, having children, and labour status. Also, we have two terms for depression and anxiety indexes, also two categorical variables for an individual's worries about his/her own health or about the health of the society. One last term shows the perception of the individual about the lockdown.

For the second part, we should use a truncated version of the count data models. The reason for truncating is that during the lockdown some people were necessarily banned to go out, and including them in the model might result in a biased estimation. Between the Poisson and the Negative Binomial, as for the overdispersion in our data (mean of 2, and variance of 4), we choose the truncated negative binomial. We should note that, as they are nested models, they can be compared with the likelihood ratio test, which also verified the dominance of the Negative Binomial for this case.

$$D_i^* = G(X_i'\beta_1 + Sad\&Dep_i\beta_2 + Anxiety_i\beta_3 + W\_ownh_i\beta_4 + W\_societyh_i\beta_5$$
$$+Perception_i\beta_6 + \epsilon_i) \ if \ D_i = 1$$

The model has the same specifications, with a count nature, and a condition of having only strictly positive values.

Additionally, as we mentioned before, if there is more than one possible explanation for zero values in the data, zero-inflated models can be powerful. Here, one part of the people does not go out as they are restricted to do so (here, for this lockdown, the restriction means not having a valid reason to go out). However, one can suspect that some people can avoid going out if they do not have an important task to do even if they are not restricted. Thus, here we create a dummy variable that takes a value of 1 if the person has a valid reason to go out, and 0 otherwise. Then, we use this variable to inflate the zero values and see how the model behaves with people who are not restricted.

For testing the usage of this model, we can do the Vuong test, which showed that the zero-inflated model did better than the basic negative binomial. However, to compare the zero-inflated model with the truncated negative binomial, the AIC, BIC, and Likelihood preferred the truncated model.

Lastly, we should also remind you that our data is mainly panel data as it includes two different periods of the survey with the same individuals. However, the periods are not two different lockdowns, but rather two different months in the same lockdowns. Thus, as expected and checked for separately, there is no difference between them for the coefficients. Thus, we are going to treat them as a whole.

5 Results

As a result, we constructed a Logit, truncated negative binomial, and zero-inflated negative binomial, where the inflation term was the variable that shows if the individual has a significant reason to go out or not. Below, in Table 1, you can see some of the noteworthy results, and the full results in Table 2.

It is obvious that being male played a big role both in the probability of going out and the number of days the individual went out, during the Covid. Even though the probability increases by only 5%, the number of days that a male individual went out is 32% more.

For the age variable, we have a significant but very low value for the probability of going out, with an insignificant and low value for the count. It is highly expected as our data mainly consists of people between the age of 20 and 55, which are considered senior and have similar reasons for going out. Thus, we can conclude that age did not play a big role in our model.

Labour term, on the other hand, shows significant effects in each of the models, which is expected as working is one of the main valid reasons for going out.

Our main variables, sadness & depression, and anxiety are very low and insignificant for the probability part. That is also an expected output, because, if an individual has a significant reason to go out, he/she will do it at least for a few days (shopping once a week, for example). Our main result comes up when we look at the effect on the number of days data. Here, they show an opposite direction of effects, where one additional level of depression causes people to go out 6% less, while a similar change in anxiety increases the number by 5.6% (note that both are significant). The size of effects can be argued, but we should note that these indexes have a range from 0 to 10. Thus, when we compare people with high depression (depression index higher than 7), and low depression (smaller than 3), we see a considerable difference in their mean value of going out.

Even though we expected a different result, worries about social health did not affect the behavior much. Only, the highest level of worry here shows a decline in the number of days, by 44% with 10% of significance. A similar scenario happens for those who worry about their health, where very worried people tend to go out 11.6% less than the non-worried people.

Lastly, commenting on the zero-inflated model, we can say that we don't get what we want. Even though, the coefficients of sadness & depression, and anxiety variables declined, the log-likelihood shows that the model did not do a good job. Here we should note that the inflating variable – dummy variable shows if the individual has a significant reason to go or not – is the only source of zero values (check Table 7 and Table 8). Because the number of zero values is similar to the number of people without a necessary reason, sticking with the truncated negative binomial gives us the most reliable result for this study.

Table 1: Specific significant results

	Double-hurdle		Zero-inflated
	Days (Binary)	Days (Count)	Days (Count)
Being male	0.058^{***}	0.315^{***}	0.218^{***}
	(0.018)	(0.061)	(0.037)
Age	0.002^{**}	0.006	0.005
	(0.001)	(0.019)	(0.012)
Labour	0.082^{***}	0.167^{*}	0.116^{**}
	(0.025)	(0.086)	(0.053)
Sadness & Depression	-0.001	-0.061^{***}	-0.040^{***}
	(0.005)	(0.019)	(0.011)
Anxiety	0.008	0.056^{***}	0.036^{***}
	(0.005)	(0.018)	(0.011)
Worry of the Own Health (3)	-0.116^{***}	0.082	-0.066
	(0.043)	(0.146)	(0.090)
Worry of the Society's Health (3)	-0.010	-0.368^{*}	-0.240^{*}
	(0.067)	(0.220)	0.134
constant	-1.623^{*}	0.559	0.908^{***}
	(0.877)	(0.452)	(0.179)
Log-Likelihood	-845.53779	-2927.6362	-3315.14
Pseudo R	0.0357	0.0119	
N	1923	1923	1923

Note: $^{*} p < 0.1$; $^{**} p < 0.05$; $^{***} p < 0.01$

Note: Marginal effects are reported for the Logit model.

6 Weak aspects of the paper

One of the weaknesses of the study is the high level of endogeneity. Firstly, there is not only the effect of depression on days of going outside the house but also the effect of the number of days leaving the house on the level of depression. So the simultaneity bias takes place. We tried to eliminate it by introducing control variables,for example, with variables that indicates if person had a compulsory reasons to go out but it significantly decreased our dataset so we decided not to use it.

Secondly, the data has two waves and individuals can duplicate but the period of all study is the 27^{th} of March to the 15^{th} of April and the quarantine regulations stayed more or less the same in this time interval. So we consider these people as different observations because they could have different levels of depression in the waves of the survey.

And finally, there could be problem of data limitations as some significant variables like level of income were not available. Both the numbers of observations and the variables could be improved to have a more reliable result. So problem of endogeneity is still relevant.

7 Improvements and Conclusion

Our study is focused on the effect of depression on going outside at the beginning of COVID-19. It is interesting itself because in a full lockdown people had more possibilities to stay home as working or studying became online and there is a lack of reasons to go outside as only the most important places are open. The data and the model captured some effects of the mood, and people's worries on the number of the days they go out. Previous research found the effect of not going out on the mood, and mixing the result, we can say that there is possibility of a vicious circle in terms of staying at home and the mood. Governments might try to be more reluctant to decide long lockdowns in the future.

At the same time it would be great to also take a look at pre- and post-covid periods. And check the hypothesis that in the beginning of the pandemic (our period) an effect is the highest.

Also the significant part of the study is the fact that regulations did not change in the period of the survey, so we can tell that the effect of the introduced rules is the same for all participants. But also it means that we cannot compare people from other countries or periods of time with different amounts of restrictions of going out. We only know that period of study is the most restricted in pandemic and analyse people behaviour in these circumstances.

8 Appendix

Table 2: Full result

	Double-hurdle		Zero-inflated
	Days (Binary)	Days (Count)	Days (Count)
Being male	0.432^{***}	0.315^{***}	0.218^{***}
	(0.139)	(0.061)	(0.037)
Age	0.109^{***}	0.006	0.005
	(0.034)	(0.019)	(0.012)
Age^2	-0.001^{***}	-0.000	-0.000
	(0.000)	(0.000)	(0.000)
Children	-0.204	-0.001^{*}	0.003
	(0.145)	(0.065)	(0.040)
Labour	0.553^{***}	0.167^{*}	0.116^{**}
	(0.158)	(0.086)	(0.053)
Sadness & Depression	-0.005	-0.061^{***}	-0.040^{***}
	(0.039)	(0.019)	(0.011)
Anxiety	0.057	0.056^{***}	0.036^{***}
	(0.038)	(0.018)	(0.011)
Worry of the Own Health (1)	-0.149	-0.060	-0.044
	(0.234)	(0.106)	(0.065)
Worry of the Own Health (2)	-0.172	-0.114	-0.082
	(0.245)	(0.114)	(0.068)
Worry of the Own Health (3)	-0.766^{***}	0.082	-0.066
	(0.286)	(0.146)	(0.090)
Worry of the Society's Health (1)	0.171	-0.026	-0.003
	(0.522)	(0.231)	(0.141)
Worry of the Society's Health (2)	0.069	-0.280	-0.182
	(0.479)	(0.216)	(0.132)
Worry of the Society's Health (3)	-0.074	-0.368^{*}	-0.240^{*}
	(0.483)	(0.220)	0.134
Perception	0.001	0.004^{*}	0.003^{**}
	(0.004)	(0.002)	(0.001)
constant	-1.623^{*}	0.559	0.908^{***}
	(0.877)	(0.452)	(0.179)

Note: $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table 3: Sex

sex	Freq.	Percent	Cum.
Female	1,173	60.31	60.31
Male	772	39.69	100
Total	1,945	100	

Table 4: Children

Children	Freq.	Percent	Cum.
No	838	43.08	43.08
Ns/Nc	28	1.44	44.52
Yes	1,079	55.48	100
Total	1,945	100	

Table 5: Worried about own health

Worried about own health	Freq.	Percent	Cum.
Ns/Nc	1	0.05	0.05
A lot	193	9.93	9.98
Few	903	46.47	56.46
No	168	8.65	65.11
Quite	678	34.89	100
Total	1,943	100	

Table 6: Worried about society health

Worried about society health	Freq.	Percent	Cum.
Ns/Nc	2	0.10	0.10
A lot	650	33.54	33.64
Few	158	8.15	41.80
No	25	1.29	43.08
Quite	1,103	56.91	100
Total	1,943	100	

Table 7: Days going out

Days going out	Freq.	Percent	Cum.
At least 1 day	1,611	82.83	82.83
Not going out	334	17.17	100
Total	1,945	100	

Table 8: Going out due to necessary reason

Necessary	**Freq.**	**Percent**	**Cum.**
Does not have necessary reason	378	19.42	19.42
Has necessary reason	1,568	80.58	100
Total	1,946	100	

References

Aguilar-Latorre, A., Oliván-Blázquez, B., Porroche-Escudero, A., Méndez-López, F., García-Gallego, V., Benedé-Azagra, B., and Magallón-Botaya, R. (2022). The impact of the COVID-19 lockdown on depression sufferers: a qualitative study from the province of Zaragoza, Spain. *BMC Public Health*, 22(1):780.

Fullana, M. A., Hidalgo-Mazzei, D., Vieta, E., and Radua, J. (2020). Coping behaviors associated with decreased anxiety and depressive symptoms during the COVID-19 pandemic and lockdown. *Journal of Affective Disorders*, 275:80–81.

Planchuelo-Gómez, , Odriozola-González, P., Irurtia, M. J., and de Luis-García, R. (2020). Longitudinal evaluation of the psychological impact of the COVID-19 crisis in Spain. *Journal of Affective Disorders*, 277:842–849.