UCD CENTRE FOR ECONOMIC RESEARCH

WORKING PAPER SERIES

2025

SRAITH PÁIPÉAR OIBRE AN IONAID

UM THAIGHDE EACNAMAÍOCHTA COBÁC

2025

What Factors Are Associated with the Decline in Young People's Mental Health During the Early Stages of the Covid Pandemic?

David Madden
University College Dublin School of Economics

WP25/04

March 2025

UCD SCHOOL OF ECONOMICS

UNIVERSITY COLLEGE DUBLIN

SCOIL NA HEACNAMAÍOCHTA COBÁC

COLÁISTE NA HOLLSCOILE BAILE ÁTHA CLIATH

BELFIELD DUBLIN 4

What Factors Are Associated with the Decline in Young People's Mental Health During the Early Stages of the Covid Pandemic?

David Madden

(University College Dublin)

March 2025

Abstract:. The Covid pandemic arrived in Ireland on February 29, 2020. In the following weeks various restrictions were introduced to stem the spread of the disease. Anxiety over the spread of the disease and over the restrictions introduced led to concerns regarding mental health. This paper uses the special Growing Up in Ireland (GUI) survey of young adults aged around 23 in December 2020 to examines the change in mental health compared to the last pre-Covid GUI survey, wave 4. In particular, it applies machine learning (ML) techniques to examine what variables are associated with the transition into depression between the two surveys. Mental health problems in wave 4 are found to play the most significant role with little effect for socioeconomic variables.

.

Keywords: Mental health; depression; lasso; Covid

JEL Codes: I14, I31.

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

What Factors Are Associated with the Decline in Young People's Mental Health During the Early Stages of the Covid Pandemic?

Introduction

The Covid 19 Pandemic officially arrived in Ireland on February 29, 2020 with the first confirmation of a positive case. Over subsequent weeks various restrictions were introduced to stem the spread of the disease (becoming collectively known as the "lockdown"). These included the closure of all educational establishments and childcare facilities, the banning of various sporting and cultural events and then on March 27, everyone, apart from providers of essential care and services, were advised to stay at home apart from essential visits (e.g. to the supermarket) and exercise within a 2km radius. There was a ban on non-essential travel and on meeting people outside the immediate household.

As Covid cases declined over the summer of 2020 there was a gradual removal of the most severe of these restrictions, but an upsurge of cases in autumn 2020 led to a reimposition of high level restrictions in October. As the second wave of Covid receded there was an easing of restrictions from early December with the opening of non-essential shops and services, including bars and restaurants, and by December 18 limited within-country travel and household visits were permitted. However, there was a significant resurgence of cases in the immediate run-up to and aftermath of Christmas and severe restrictions were again imposed in January 2021.

The various lockdowns were successful in limiting the spread of Covid but these measures were not without their own costs. The most severe restrictions inevitably led to a reduction in economic activity and the reduction in human contact and the hardship imposed by social distancing raised concerns about possible adverse mental health effects. We investigate the latter phenomenon in this paper. The landmark *Growing Up in Ireland* (GUI) longitudinal study carried out online surveys, including questions on mental health, for both the infant (2008) and child (1998) cohort for most of the month of December 2020 (the 2008 cohort survey began on December 4 and the 1998 cohort survey began on December 11, both surveys ending at the end of December). The 1998 cohort was aged between 22 and 23 at the time of this survey. The 2008 cohort was aged 12 at the time of the survey and their primary care givers (PCGs, in almost all cases the biological mothers) were also surveyed. This group ranged in age from 33 to 55. No data were collected for the PCGs of the older cohort.

GUI thus provides a snapshot of mental health for these groups at a time when restrictions were still relatively tight and had been ongoing for nearly nine months. Critically, since this is a longitudinal study we also have similar information *for the same people* for a period *before* Covid, for the 2008 cohort between June 2017 and February 2018, and for the 1998 cohort between August 2018 and June 2019 (the periods over which fieldwork was carried out). It is the 1998 cohort which is the subject of this paper.

While the 1998 cohort is by no means representative of the whole nation, it is representative of those people born in that specific period, and what may be lacking in national representativeness is compensated for by having the same measure, for the same people before and during the pandemic. In a previous paper (Madden, 2024) we documented the change in mental health for both GUI cohorts, showing how overall a deterioration was experienced. In that paper we also checked for the presence of a socioeconomic gradient to the deterioration in mental health and found that, with the exception of the PCGs of the study children for the 2008 cohort, there was no statistically significant evidence of such a gradient. This is somewhat unusual in that many morbidities, including mental health, exhibit some form of socioeconomic gradient (Reiss, 2013, Devenish et al 2017). In this paper we investigate this issue further, taking a different approach to our earlier work. More specifically, we apply machine learning (ML) techniques to our data to see which of the many variables in GUI were the best predictor for becoming depressed between wave 4 and the Covid survey. We confine our analysis to the young adults of the 1998 Cohort i.e. those aged 22-23 at the time of the GUI Covid survey, as for this group we have a consistent measure of mental health for both before and during Covid.

Mental Health and the Covid Pandemic

A number of studies have examined the impact of Covid restriction on various measures of mental health or well-being. Aknin et al (2022) found that the restrictions introduced at the initial stages of Covid (around March 2020) impacted upon subjective well-being and psychological distress, but that much of this impact had abated by June 2020. Banks and Xu (2020) found substantial population level effects in a survey carried out in April 2020, with particularly marked effects upon young adults and women. Davillas and Jones (2021) also found substantial effects early in the pandemic but that much of this had gone by July 2020. Hajek et al (2022a, 2022b) investigated anxiety and depression across seven European countries during the later stages of Covid (from November 2020 to September 2021) and found highest rates for young people, aged 18-29. Tofolutti et al (2022) looked at the impact of Covid policies on mental well-being for 28 European countries from April 2020 to March 2021. They found that restrictions on international travel, private gatherings and contact tracing were associated with reductions in mental well-being of up to 4 per cent, with greater effects for females and those living with younger children. However, they do not have information on mental well-being before Covid.

Kauhanen et al (2023) carried out a meta analysis of studies which examined changes in mental health arising during the pandemic. They stressed the importance of having longitudinal data which follow the same set of people before and during the pandemic and also include the same measure of mental health. They found relatively few studies satisfying such criteria but those studies which were identified showed deteriorating mental health for a variety of measures used, such as depression and anxiety.

Orgad (2024) reviewed the association between lockdowns and a variety of measures of mental health, again finding that stricter lockdowns were associated with worse mental health outcome although it was not clear whether the poor mental health outcomes arose on account of the lockdowns, or the underlying Covid situation which may have prompted the lockdowns.

In terms of studies for Ireland, Hyland et al (2021) found no evidence of an increase in mental health problems for Irish adults during the first year of the pandemic. It is noticeable however that they had no measures of mental health for their sample before the pandemic, so while mental health may not have deteriorated subsequent to the arrival of the pandemic, their analysis cannot tell if the pandemic caused an immediate deterioration in mental health.

Smyth and Murray (2022) used the GUI Covid module to investigate mental health for the 2008 cohort. However, since they did not have a consistent measure of mental health pre and post Covid their study was unable to explicitly examine how mental health changed with the pandemic. They analysed the association between the level of mental health in December 2020 and various factors such as family financial and education resources and restrictions on social activities. They found that lower mental health for the 2008 cohort was associated with a fall in family income arising from the pandemic (as opposed to family income pre-pandemic) and also with lower educational resources in terms of access to a computer and/or a quiet place to study.

The advantage provided by our study is that of a larger dataset than has been the case for other studies for Ireland (excepting Smyth and Murray) and critically the availability of the same measure of mental health for the same people both before and during Covid. This is in contrast to other studies which have only followed mental health after the onset of the pandemic. The richness of our dataset allows us to investigate a wide range of variables, as measured in wave 4, which could be associated with transitions into adverse mental health between the last pre-Covid survey and the Covid survey, by adopting a machine learning approach to select such variables.

Data and statistical analysis

Data

Our data consist of a cohort of young people born in the period November 1997-October 1998 (Williams et al, 2009). The specific GUI data we analyse is the last available (pre-Covid) wave of this cohort, wave 4 (collected in 2018/19) and the Covid survey which was sent out in December 2020.

The last pre-Covid wave of this cohort consisted of 5190 young adults and the original sample frame was the national primary school system, with 910 randomly selected schools participating in the study. The following exclusions were placed on the data: a balanced panel only was used i.e. observations who responded to both the Covid survey and the last pre-Covid survey. In addition, observations where the questions on mental health were not answered were also excluded. Following these exclusions, this left 1950 observations (1243 females and 707 males).[1]

The measure of mental health used is the Center for Epidemiologic Studies Depression Scale (CES-D, Melchior et al 1993). The original version of the CES-D scale had 20 items and

[1] It is worth pointing out that of the 5190 observations in wave 4, these 1950 observations showed measured mental health in wave 4 which was slightly *better* than that of the remain 3240 observations i.e. those who responded to the Covid survey were not the less healthy of the wave 4 sample (details available on request from the suthor).

has been used extensively across the world and has featured in many published journal articles. There are also shorter versions of the measure which take less time to administer but are still regarded as reliable measures of mental health. One of these is the CES-D8, with eight items, and this is the version which is measured in GUI.

The CES-D8 measure consists of eight statements regarding how the respondent was feeling in the past week (e.g. "I felt depressed", "I felt fearful" etc). The respondent then indicates whether they experienced this feeling rarely/none of the time, some or a little of the time, occasionally or a moderate amount of the time or most or all of the time. Answers are coded 0, 1, 2 or 3 respectively, so that the minimum score possible is 0 and the maximum is 24. Higher scores indicate worse mental health and a score at or above 7 is regarded as indicating depression (Devins et al, 1988). The data in GUI is truncated at 13 (i.e. all CES-D8 scores greater than or equal to 13 are coded as 13).

We wish to examine the factors lying behind the observed deterioration in mental health and so the question arises as to how to measure this "deterioration". One possibility is to simply look at the change in the CES-D8 measure. There are some problems with this, however. First, it effectively treats the CES-D8 as though it were cardinal, as opposed to an ordered-categorical, scale. Secondly, it would not distinguish between people who register a given increase of, say 4, in the CES-D8, some of whom cross the key threshold of 7, and others who do not. Thirdly, recorded CES-D8 values in GUI are truncated at 13, so in some cases while a person's CES-D8 has increased, if it hits the 13 threshold, the full extent of the change will not be recorded.

We thus choose to measure the deterioration in mental health with a binary variable as those people who become depressed between wave 4 of GUI and the Covid survey viz. those who had a CES-D8 value below 7 in wave 4, but have a value of 7 or above in the Covid survey. We label this group as "became depressed" and so our sample for analysis is thus those who potentially could have become depressed between wave 4 and the Covid survey i.e. they had CES-D8 values below 7 in wave 4. We thus exclude 538 observations who were above this threshold pre-Covid, leaving a sample of 1412 (556 male and 856 female) of whom 556 cross the key threshold of becoming depressed between wave 4 of GUI and the Covid survey.[2] Note

[2] Although attrition from GUI is not random, the machine learning packages we employ in the paper do not support the use of sampling weights. However note that our analysis will include as control variables those covariates upon which sampling weights are typically based.

that of the 538 excluded who were depressed in wave 4, 128 of them made the transition in the other direction and are thus classified as "non-depressed" in the Covid survey. While examining the factors associated with this transition would be an interesting exercise, they are not the subject of this paper.

Note that given the nature of the GUI surveys, we only observe the sample at snapshots in time and thus our measure effectively attributes *all* changes in mental health between wave 4 and the Covid survey to the effect of the pandemic. This is problematic, given that mental health problems are generally held to exhibit an inverse U relationship with age, and so we might expect some deterioration in mental health even in the absence of Covid. Mental health can also exhibit seasonal patterns which we are unable to include. Ideally, we would like to compare mental health in December 2020 with a counterfactual of what mental health would have been in December 2020 in the absence of Covid. However, given the nature of our data, we are unable to make this comparison, but data from elsewhere for young adults suggests that a substantial proportion of the decline in mental health in 2020 is due to Covid. For example, Banks and Xu (2020) in their analysis of the deterioration in mental health for 16-24 year olds between Jan-May 2019 and April 2020 in the UK, estimate that about three-quarters of the change is attributable to Covid.

Table 1a shows the distribution of the sample by depressed/non-depressed status before and after Covid. Table 1b shows the transitions into depression between wave 4 before Covid and the Covid wave. 386/856 females became depressed, about 45 per cent. For males, the number was 170/556, around 30 per cent.

Statistical Analysis

The purpose of this paper is to try to find which, of the many variables present in GUI, are the best predictors of transitioning into depression between wave 4 and the Covid survey. As described by Mullaithanan and Spiess (2017): "[machine learning]… revolves around the problem of prediction: produce predictions of *y* from *x*. The appeal of machine learning is that it manages to uncover generalizable patterns.". This distinguishes it from traditional econometric practice which seeks to obtain estimates of parameters which underly the relationship between *y* and *x*. Although our main focus in this paper is to find best predictors, we do also recover parameter estimates and we will discuss those briefly.

We applied machine learning using the *lasso* function in Stata. Although generally used as a term in its own right these days, lasso is an acronym for "least absolute shrinkage and selection operator". Lasso is primarily used as a means of selecting covariates for a model and avoiding the problem of overfitting. Given an outcome variable *y* and a vector of potential covariates *X*, analysts may typically choose a subset of *X* to provide a prediction of the outcome *y*. While the analyst may choose certain variables from *X* on the basis of theory or intuition, given a sufficiently large set of *X*, there may be hundreds of candidate variables.

Adding extra variables to a model will always improve the "fit" of the model in the sense of increasing the R^2, but this comes at the expense of over-fitting, whereby a model may give a very close fit for the sample upon which it is estimated, but give a very poor fit outside this sample. Lasso regression is one way of addressing this as it applies a penalty as more and more variables from *X* are included in the model. More specifically, suppose our regression model is of the linear form

$$y = \beta_1 \boldsymbol{x_1} + \beta_2 \boldsymbol{x_2} + \cdots \beta_p \boldsymbol{x_p} + \epsilon$$

where y is the outcome variable and the $\boldsymbol{x_i}$ are covariates with associated coefficients β_i, with error term ϵ. Least squares estimation will find the values of β_i which minimize

$$\frac{1}{N}(\boldsymbol{y} - \boldsymbol{X\beta}')'(\boldsymbol{y} - \boldsymbol{X\beta}')$$

where $\boldsymbol{X}$ is the nxp matrix of covariates and $\boldsymbol{\beta}$ is the px1 vector of coefficients i.e. the sum of squared errors. Lasso estimation however will minimize

$$\frac{1}{N}(\boldsymbol{y} - \boldsymbol{X\beta}')'(\boldsymbol{y} - \boldsymbol{X\beta}') + \lambda \sum_{j=1}^{p} |\beta_j|.$$

The extra term $\lambda \sum_{j=1}^{p} |\beta_j|$ is a penalty term which increases in value the more complex the model becomes i.e. the more extra variables are added. It thus penalizes over-fitting. The lasso regression minimizes the above expression for given values of λ, and then chooses the optimal λ, λ^*, based on a number of possible criteria. The three lasso approaches we use to choose λ^* are: CV (cross validation) lasso, adaptive lasso and BIC lasso. We run these versions using both a linear and logit model, given that our dependent variable (becoming depressed between wave 4 and the Covid wave) is binary. We now explain each of these three approaches in more detail.

With the CV approach, a grid of a specified number of λs is set, starting off with the minimum value of λ which gives no non-zero coefficients. As λ becomes smaller, variables with non-zero coefficients are added (and occasionally some are removed) until the minimized value of the CV function is attained. The CV function is obtained by dividing the sample into a number of different folds (in this application 10 were used). One fold is chosen, and then a regression is fit on the other nine folds using the variables in the model for that λ. Then, using these new coefficient estimates, a prediction is computed for the data of the selected fold. The mean squared error (MSE) of the prediction is computed and this process is then repeated for the other nine folds. The 10 MSEs are then averaged to give the value of the CV function. This process stops when a minimum of the CV function is obtained and the corresponding value of λ is λ^*.

With the adaptive approach, more than one, usually two, lassos are run. In the first, a value of λ^* is obtained, and penalty weights are constructed from the coefficient estimates. These weights are then used in a second lasso, and another λ^* is obtained and usually this is the one chosen, as λ^* typically changes very little after the second lasso. Thus in the ultimate lasso, usually the second, the penalty term is now $\lambda \sum_{j=1}^{p} \widehat{w_j} |\beta_j|$, where the $\widehat{w_j}$ are the penalty weights.

The final approach used is the BIC approach where λ^* is chosen to minimise the Bayesian Information Criterion.

For all of these models we use 70 per cent of the sample as our "training data" and we then apply the chosen model to the remaining 30 per cent "testing" data to check model fit. We measure fit in a number of ways. First, we report goodness of fit measures, R^2 for linear models and deviance ratios for the logit models. We also calculate the area under the receiver operating characteristic curve for the forecast values obtained when applying the chosen model to the testing data. The area under the curve (AUC) represents the probability that the model, if given a randomly chosen positive (became depressed) and negative (did not become depressed) observation from the testing data, will rank the positive higher than the negative. An AUC of 1.0 represents a "perfect" model, in that it will always perfectly predict if an observation will be positive or negative. Perhaps counterintuitively, an AUC of zero is also a perfect predictor, in that it always perfectly predicts the “wrong” binary state, it is a perfect counter-predictor. From a forecasting perspective the “worst” value for an AUC to take is 0.5, since then the chances of the model prediction being correct is 50-50, and a coin toss would perform just as well. While there are no precise guidelines, in terms of binary classification, an AUC of 0.7-0.8 is considered acceptable (Mandrekar, 2010).

As a visual aid to assessing goodness of fit, we also present calibration belts (Nattino et al., 2017). These are graphs which show the goodness of fit of binary outcome models, by observing the relationship between probabilities estimated by the model in question, p_i and “true” probabilities $P(Y_i = 1)$, where the true probabilities are obtained from a polynomial logistic regression of the form $logit\{P(Y_i = 1)\} = \gamma_0 + \gamma_1 logit(p_i) + \cdots + \gamma_m \{logit(p_i)\}^m$. and also include an associated likelihood ratio test on how well the estimated probabilities conform to the “true” probabilities, which is essentially a test of the hypothesis $H_0: (\gamma_0, \gamma_1, \gamma_2, \dots, \gamma_m) = (0, 1, 0, \dots, 0)$. The optimal degree of m is obtained by starting off with a low value and then carrying out a sequence of likelihood ratio tests to forwardly identify m (see Nattino et al, 2016). The 45 degree line (shown in red in the figures below) indicates an exact relationship between estimated and true probabilities and the confidence band around the curve i.e. the calibration belt, reflects the statistical uncertainty about the estimate of the curve. The accompanying table shows whether the calibration belt always includes the 45 degree line, and also the regions where it does not. The p-value reported is that for the null hypothesis H_0 above.

Results

We now present results from the lasso analysis for the linear and logit cases. Note that the analysis is only applied to our sample in wave 4, as we are investigating which wave 4 variables best predict the transition into depression. In all cases the variables were standardised and for the moment, we do not discuss the relative sizes of the coefficients as we are primarily engaged in a prediction and variable selection exercise here i.e. what variables are the best predictors of becoming depressed between wave 4 of GUI and the Covid wave. Tables 2a and 2b list the variables chosen by Lasso while figures 1a to 5c provide the associated graphs for the areas under the ROC curves and the calibration belts.

We first look at the various measures of goodness of fit for the out of sample, or testing, dataset to see if any particular model stands out. In terms of R^2 or deviance ratios, essentially how the model performs in comparison to a null model which included no covariates, the preferred model is lasso with BIC with values of 0.0483 and 0.0309 respectively. These are very low values however, suggesting that the model with covariates only explains 3-4 per cent of variation in the dependent variable outside of sample. An alternative measure of goodness of fit, and arguably more appropriate in the case where we are trying to classify observations into one of two classes, is the area under the ROC curve. In this case the CV folds lasso for both linear and logit cases performs best, although its values of around 0.68 would still be regarded as a poor to weak prediction model. This model also has the highest p-value from the calibration belt test.

The CV folds model typically selects more variables than either the adaptive or BIC models and this is the case here, though it selects as many variables as the adaptive model for the linear case and only two more variables for the logit case. The BIC model is easily the most parsimonious model, only selecting two variables, the wave 4 CES-D8 score and a measure of self-esteem.

Overall it seems fair to say that none of the chosen models perform particularly well in terms of prediction, but the lasso CV folds model is the best of a weak lot. However, there is quite significant overlap between the variables chosen by the various lasso. We now see which variables rank highest in terms of their contribution to the prediction. Tables 4a and 4b present the five most important variables (in terms of the absolute values of their standardised coefficients) for each model. There is a fair degree of agreement across the models. In all cases, the value of CES-D in wave 4 is the variable with the largest standardised coefficient. Also featuring prominently is the Rosenborg Self-Esteem scale, where higher values reflect greater self-esteem. Self-Assessed health of the primary caregiver also ranks highly, and in this case, the sub-category “good” is the critical variable. It must be noted that despite the label “good”, parents who choose this value are in the lower quartile of self-assessed health, as there are two higher categories, “very good” and “excellent”. Hence, this variable is really reflecting a situation where having a PCG in the lowest quartile of self-assessed health is a positive predictor of becoming depressed.

Other variables which rank highly are lack of private medical insurance (which may partly capture a socioeconomic gradient) and not contacting an out-of-hours GP. Since this variable has a negative coefficient, it indicates that consulting a GP outside of normal hours (presumably for some form of emergency) is a positive predictor of becoming depressed between wave 4 and the Covid survey.

Overall, as was found in Madden (2024) there seems to be little role for the classic measures of socioeconomic status, such as parental education, income or class. If any pattern can be identified, it is that measures in wave 4 which indicate fragility of mental health, such as values of CES-D8 which are high but not sufficiently high to indicate depression, and low values of self-esteem, or use of an emergency GP seem to be the best predictors of transition into depression in the Covid survey. The importance of CES-D8 in wave 4 in some ways could be regarded as mechanistic, since anyone with a high value of CES-D8 (but below 7) in wave 4 has “less far” to travel to cross the key threshold of 7. However, the importance of other measures of mental health such as self esteem and emotional stability suggests that its importance is not just mechanistic.

Strengths and Limitations

The strength of this paper is the availability of high quality longitudinal data and, as emphasised by Kauhanen et al (2023) in their meta analysis, having the same measure of mental health before and during Covid for the same group of people. The wide range of variables in GUI means that lasso analysis can be applied to a rich set of factors.

Another strength of the paper is the atheoretical nature of machine learning. The analysis simply allows the data to identify what factors are found to exhibit an association with becoming depressed, without any prior restrictions being imposed.

The principal limitation of the paper is that the pre-Covid measure of mental health was collected between August 2018 and June 2019 and it is possible that some observations may have transitioned into depression after this period but *before* the arrival of Covid and hence these transitions cannot be regarded as a Covid effect. However, field work for the pre-Covid survey was carried out as late as June 2019, so the interval before the arrival of the pandemic is quite short.

The second limitation of the paper is that we are using lasso analysis as a method for variable selection for prediction. As discussed above the analysis is atheoretical but that implies the absence of structure with which to analyse the results, and hence renders the interpretation of the coefficients more difficult.

Conclusion

This paper applies machine learning techniques to investigate what variables are associated with a decline in mental health during the early stages of the Covid pandemic for a sample of young Irish adults. Three different lasso models were applied to the data and overall the model fit was fairly poor. However, there was considerable agreement across the different models as to which variables were associated with a transition into depression between the last pre-Covid GUI survey and the Covid survey. These variables all stress the role of fragile mental health before Covid, with higher risk for young people who had values of CES-D8 just below the depression threshold and also those who exhibited low self-esteem. To the extent that results from this paper have relevance for future pandemics, should severe restrictions similar to the lockdowns of 2020/21 be introduced, then people with existing mental health problems appear to be most at risk and supports should be targeted in that direction.

Table 1a: Numbers above and below CES-D8 value of 7

	CES-D8<7		CES-D8≥7		Total
	Female	Male	Female	Male	
Wave 4	856	556	387	151	**1950**
Covid Survey	552	432	691	275	**1950**

Table 1b: Transition into Depression between wave 4 of GUI and Covid wave

CES-D8<7 in wave 4		CES-D* <7 in wave 4 & CES-D8≥7 in Covid Wave	
Female	Male	Female	Male
856	556	386	170

Table 2a: Variables Chosen by Lasso Analysis (linear)

	Lasso (CV)	Lasso (Adaptive)	Lasso (BIC)
2nd level education did not help self-confidence	x	x	
2nd level education helped appreciate art	x	x	
Able to save on regular basis	x	x	
CES-D8 wave 4	x	x	x
Concerned about gender equality	x	x	
Concerned about terrorism	x	x	
Confience in media	x	x	
Consulted out of hours GP	x	x	
Consulted psychiatrist	x	x	
Deliberate vandalism to property	x	x	
Difficulty sleeping	x	x	
Emotional Stability	x	x	
Family, friends in area	x	x	
Great difficulty making ends meet	x	x	
Home/garden in bad condition	x	x	
Job allows you to be creative	x	x	
No competency in caring for others	x	x	
Not covered by private health insurance	x	x	
PCG in good health	x	x	
Posted online about politics	x	x	
Proactive	x	x	
Required attention after assault	x	x	
Satisfied with school programme	x	x	
Self-esteem	x	x	x
Semi-skilled, unskilled class	x	x	
Time spent online weekday	x	x	
Time spent online weekend	x	x	
Victim of crime	x	x	
Area under ROC Curve	0.681	0.647	0.642
R^2 (training sample)	0.1545	0.1607	0.1378
R^2 (testing sample)	0.006	0.006	0.0482
P value from calibration test	0.176		0.013

Table 2b: Variables Chosen by Lasso Analysis (logit)

	Lasso (CV)	Lasso (Adaptive)	Lasso (BIC)
2nd level education did not help self-confidence	x	x	
2nd level education helped appreciate art	x	x	
Able to save on regular basis	x	x	
CES-D8 wave 4	x	x	x
Concerned about gender equality	x	x	
Consulted out of hours GP	x	x	
Deliberate vandalism to property	x	x	
Difficulty sleeping	x	x	
Emotional stability	x	x	
Great difficulty making ends meet	x	x	
Home/garden in bad condition	x	x	
Job allows you to be creative	x	x	
No competency in caring for others	x	x	
No private med insurance	x	x	
PCG in good health	x	x	
Posted online re politics	x	x	
Proactive	x	x	
Required attention after assault	x		
Satisfied with school programme	x	x	
Self esteem score wave 4	x	x	x
Time spent online weekday	x	x	
Time spent online weekend	x		
Victim of crime	x	x	
Area under ROC Curve	0.673	0.663	0.642
Deviance Ratio (training sample)	0.0699	0.1458	0.1083
Deviance Ratio (testing sample)	-0.0727	-0.0347	0.0309
P value from calibration test	0.051	0.001	0.051

Table 4a: Top predicting variables from linear models

CV Folds Lasso		Adaptive Lasso		BIC Lasso	
Variable	**Coefficient**	**Variable**	**Coefficient**	**Variable**	**Coefficient**
CESD Wave 4	0.096	CESD Wave 4	0.108	CESD Wave 4	0.076
Self-Esteem	-0.041	Competent in caring for others	-0.078	Self-Esteem	-0.019
Did not consult out of hours GP	-0.031	Primary Care Giver in good health	0.057		
Primary Care Giver in good health	0.025	Did not consult out of hours GP	-0.056		
Not covered by private health insurance	0.022	Damage to property in your locality very common	0.055		

Table 4b: Top predicting variables from logit models

CV Folds Lasso		Adaptive Lasso		BIC Lasso	
Variable	**Coefficient**	**Variable**	**Coefficient**	**Variable**	**Coefficient**
CESD Wave 4	0.419	CESD Wave 4	0.539	CESD Wave 4	0.321
Self-Esteem	-0.193	Competent in caring for others	-0.395	Self-Esteem	-0.083
Did not consult out of hours GP	-0.131	Self-Esteem	-0.367		
Primary Care Giver in good health	0.107	Great difficulty making ends meet	0.299		
Not covered by private health insurance	0.098	Primary Care Giver in good health	0.299		

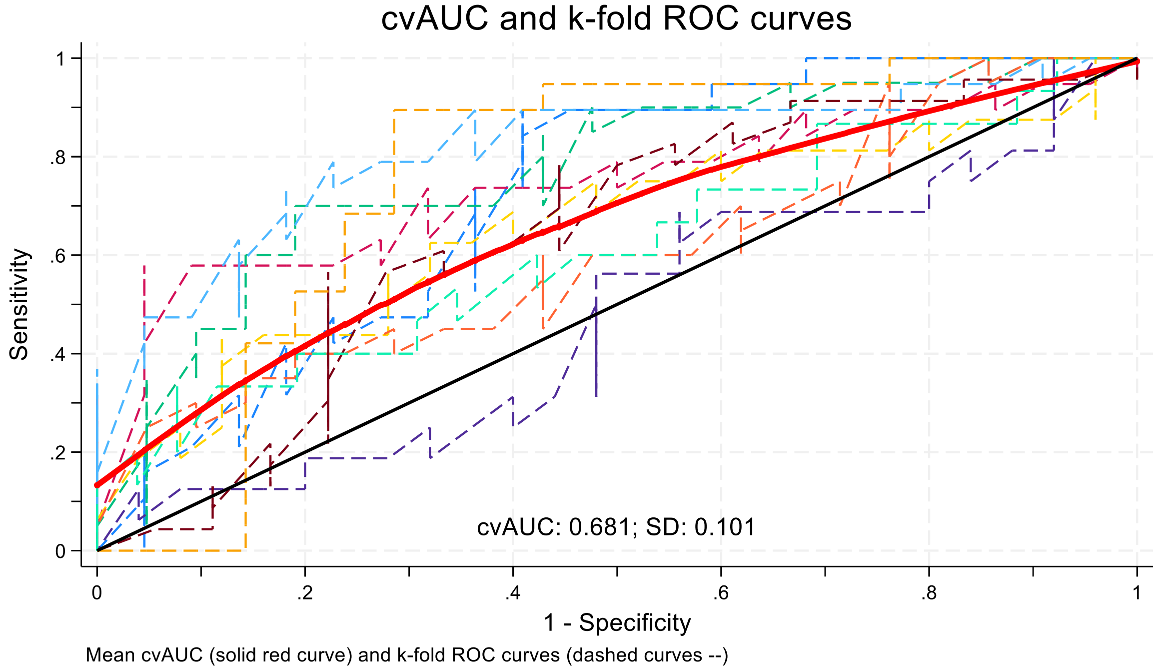

Figure 1a: Area under ROC curve, CV Folds Lasso (linear)

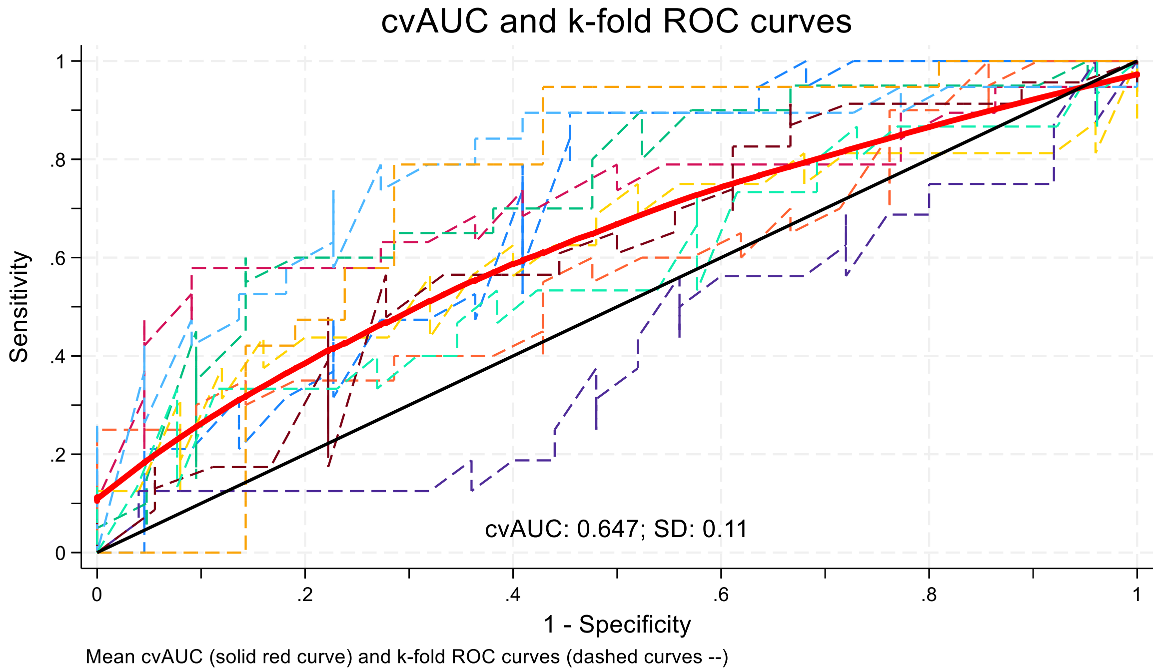

Figure 1b: Area under ROC curve, Adaptive Lasso (linear)

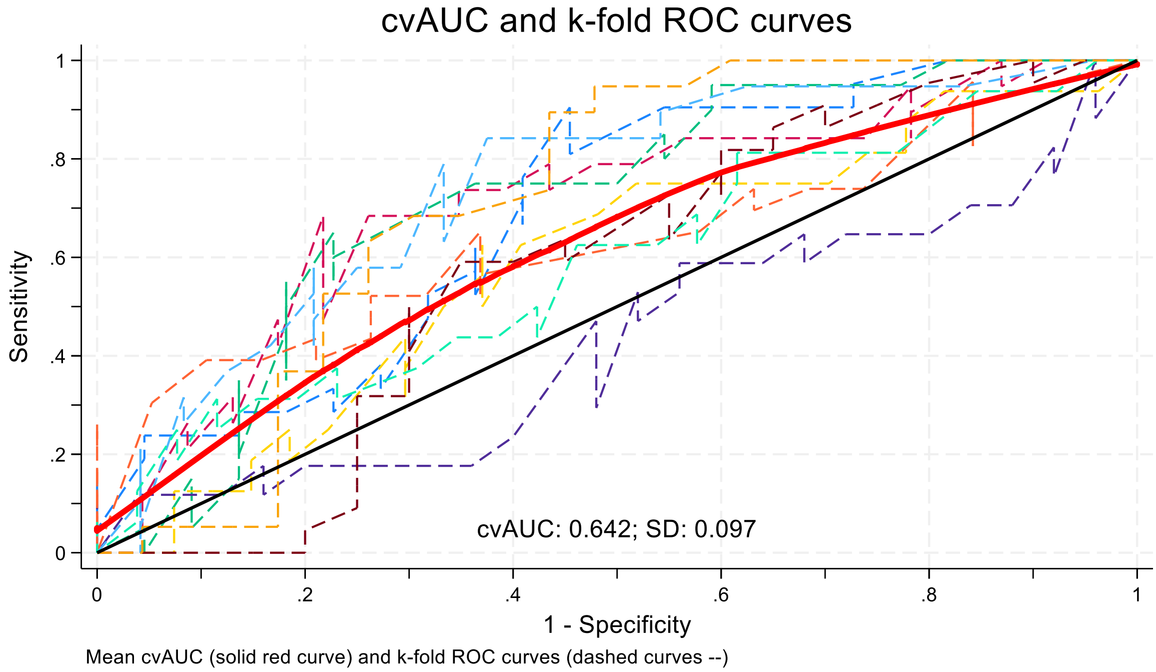

Figure 1c: Area under AUC curve, BIC Lasso (linear)

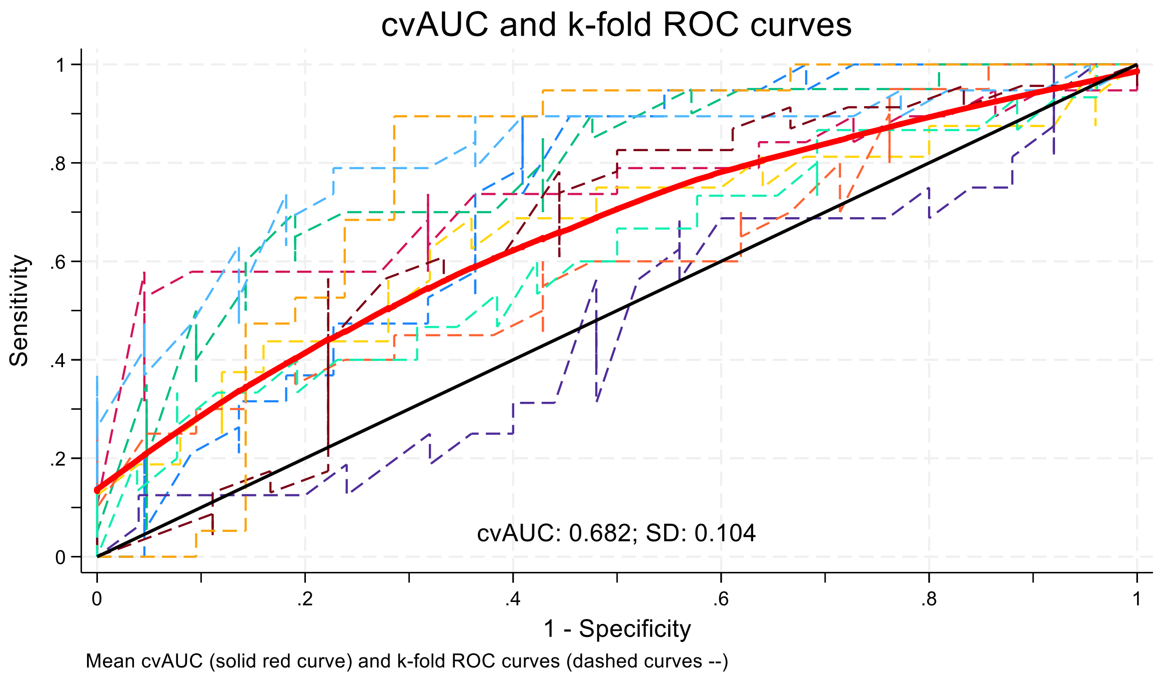

Figure 2a: Area under ROC curve, CV Lasso (logit)

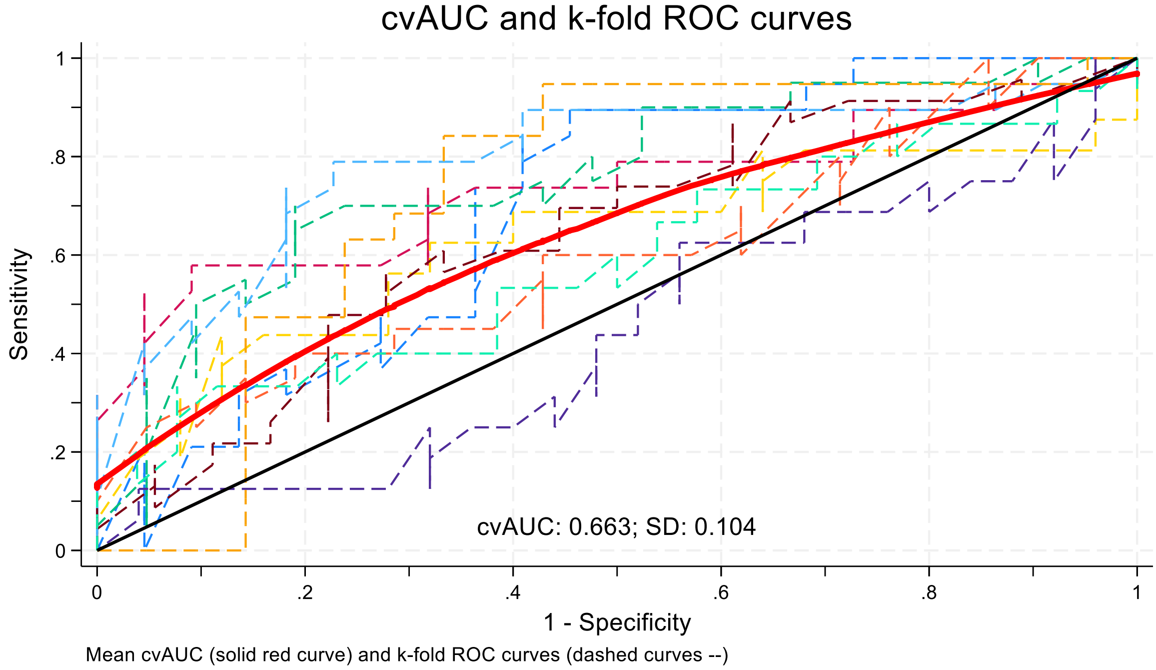

Figure 2b: Area under ROC curve, Adaptive Lasso (logit)

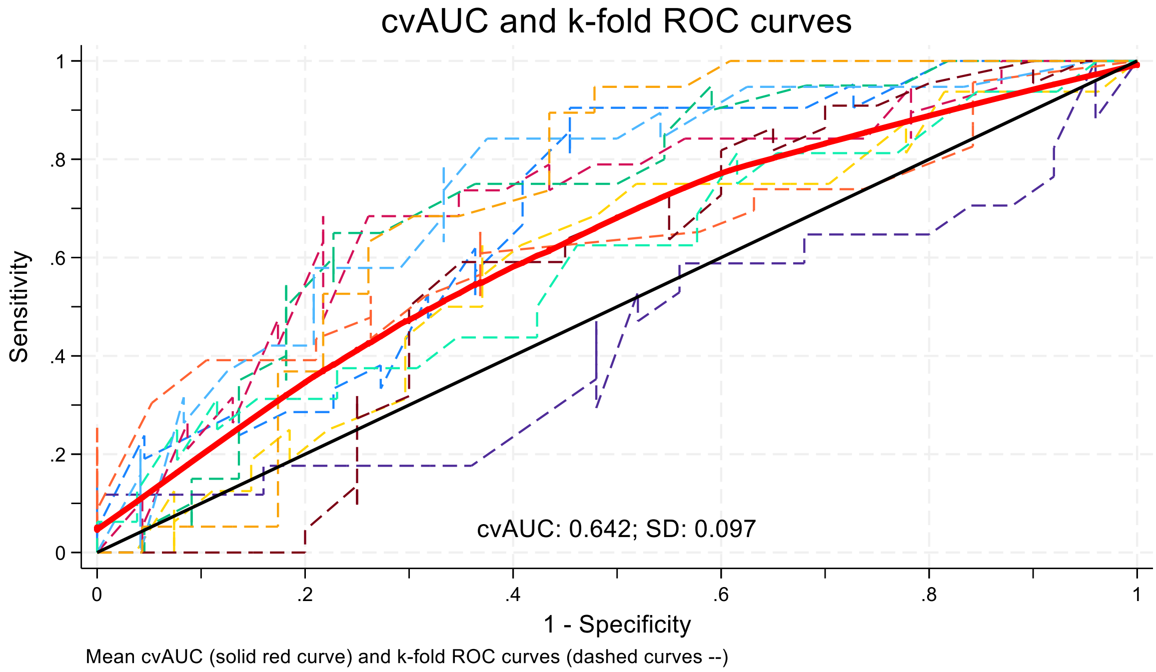

Figure 2c: Area under ROC curve, BIC Lasso (logit)

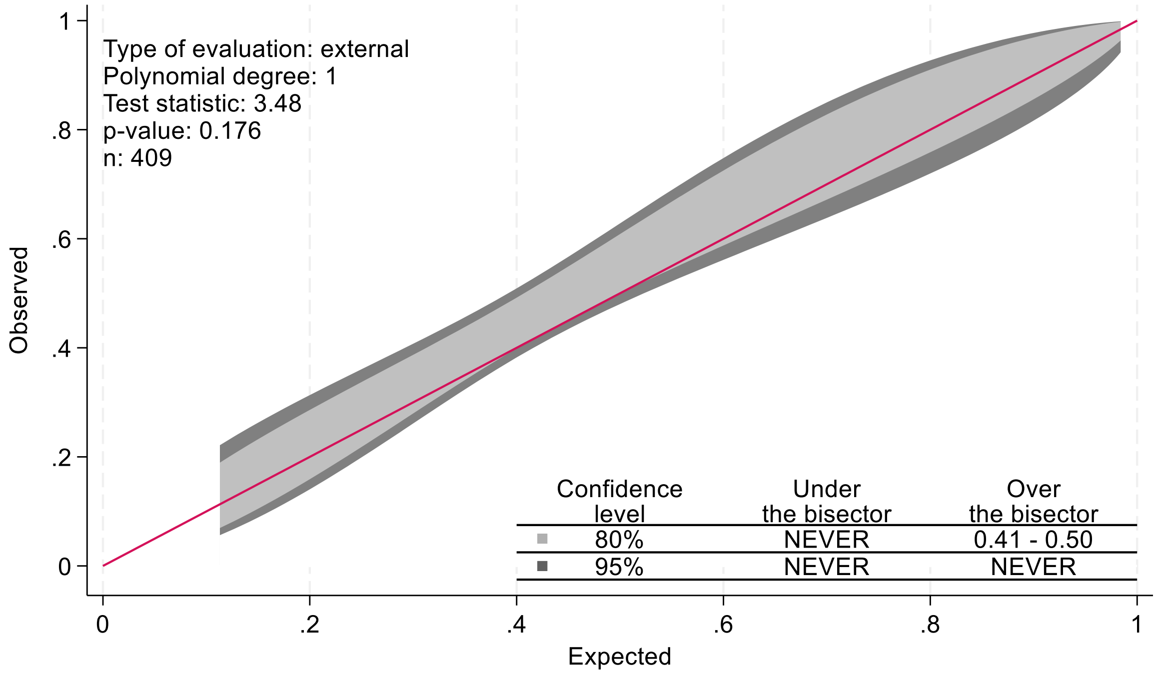

Figure 4a: Calibration Belt, CV Folds Lasso (linear)

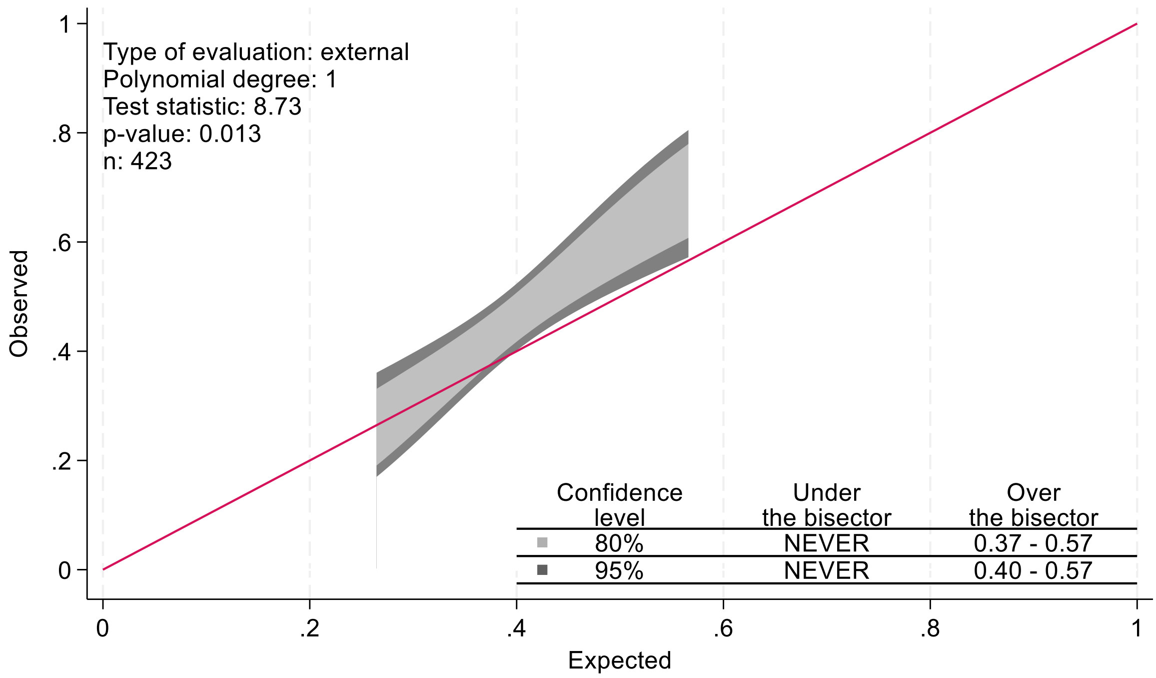

Figure 4b: Calibration belt, BIC Lasso, linear

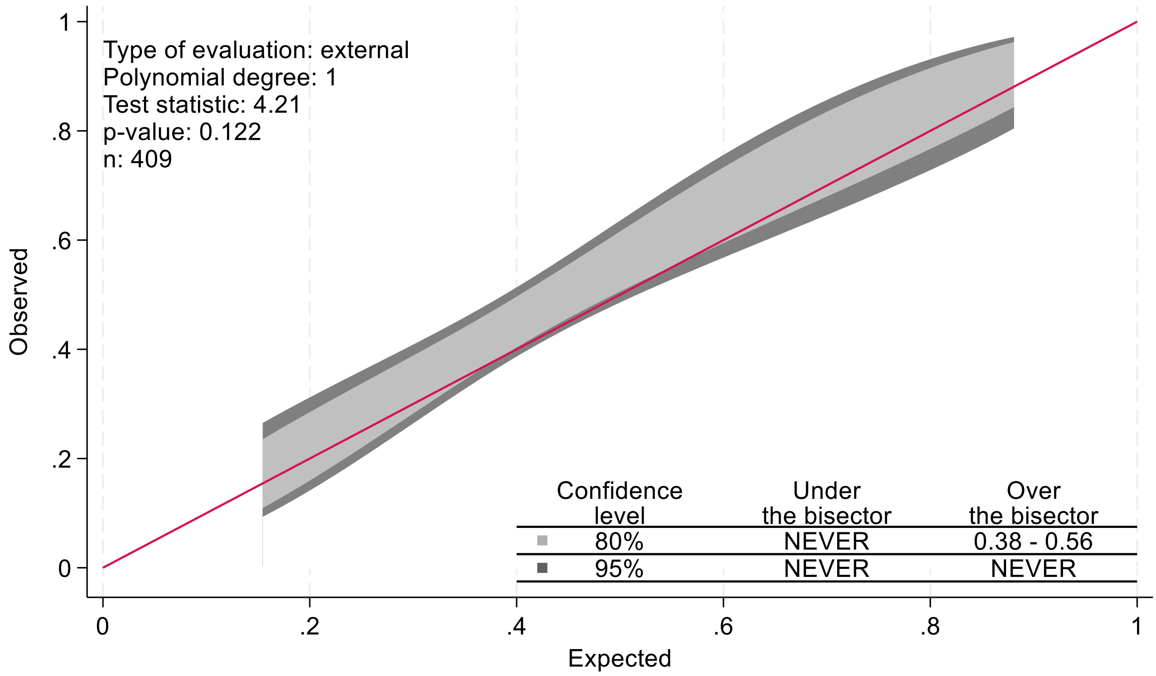

Figure 5a: Calibration Belt, CV Folds Lasso (logit)

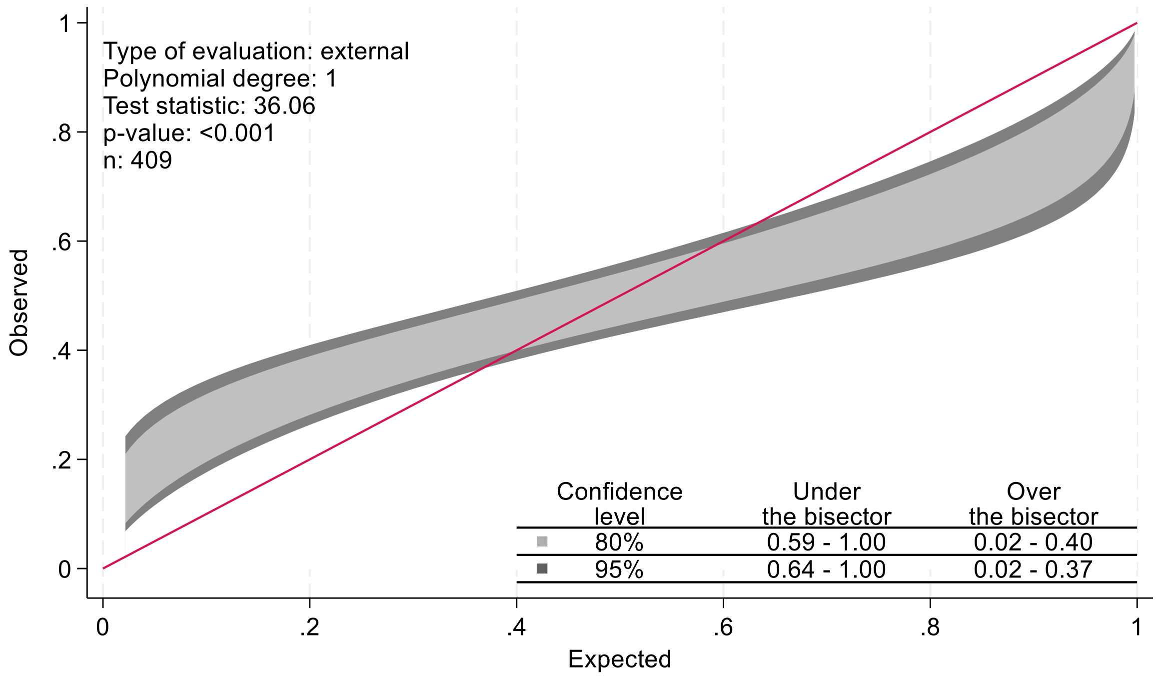

Figure 5b: Calibration Belt, Adaptive Lasso (logit)

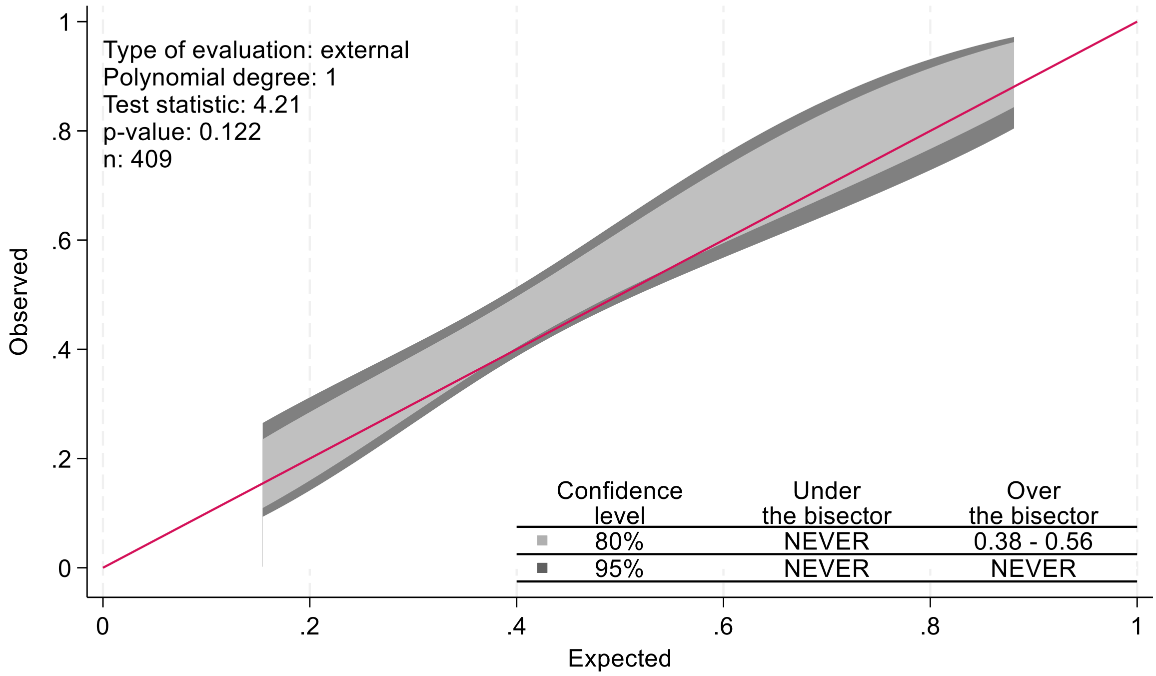

Figure 5c: Calibration Belt, BIC Lasso (logit)

References

Aknin LB, De Neve JE, Dunn EW, Fancourt DE, Goldberg E, Helliwell JF, Jones SP, Karam E, Layard R, Lyubomirsky S, Rzepa A. Mental health during the first year of the COVID-19 pandemic: A review and recommendations for moving forward. *Perspectives on Psychological science*. 2022 Jul;17(4):915-36.

Banks J, Xu X. The mental health effects of the first two months of lockdown during the COVID-19 pandemic in the UK. *Fiscal Studies*. 2020 Sep;41(3):685-708.

Davillas A, Jones AM. The first wave of the COVID-19 pandemic and its impact on socioeconomic inequality in psychological distress in the UK. *Health Economics*. 2021 Jul;30(7):1668-83.

Devenish, Bethany, Merrilyn Hooley, and David Mellor. "The pathways between socioeconomic status and adolescent outcomes: A systematic review." *American journal of community psychology* 59, no. 1-2 (2017): 219-238.

Devins GM, Orme CM, Costello CG, Binik YM, Frizzell B, Stam HJ, Pullin WM. Measuring depressive symptoms in illness populations: Psychometric properties of the Center for Epidemiologic Studies Depression (CES-D) scale. *Psychology and Health*. 1988 Apr 1;2(2):139-56.

Hajek A, Sabat I, Neumann-Böhme S, Schreyögg J, Barros PP, Stargardt T, König HH. Prevalence and determinants of probable depression and anxiety during the COVID-19 pandemic in seven countries: Longitudinal evidence from the European COvid Survey (ECOS). *Journal of affective disorders*. 2022a Feb 15;299:517-24.

Hajek A, Neumann-Böhme S, Sabat I, Torbica A, Schreyögg J, Barros PP, Stargardt T, König HH. Depression and anxiety in later COVID-19 waves across Europe: New evidence from the European COvid Survey (ECOS). *Psychiatry Research*. 2022b Nov 1;317:114902.

Hyland P, Vallières F, McBride O, Murphy J, Shevlin M, Bentall RP, Butter S, Hartman TK, Karatzias T, MacLachlan M, Maguire R. Mental health of adults in Ireland during the first year of the COVID-19 pandemic: Results from a nationally representative, longitudinal study. *Psychological Medicine*. 2021 Oct 11:1-3.

Kauhanen, Laura, Wan Mohd Azam Wan Mohd Yunus, Lotta Lempinen, Kirsi Peltonen, David Gyllenberg, Kaisa Mishina, Sonja Gilbert, Kalpana Bastola, June SL Brown, and Andre Sourander. "A systematic review of the mental health changes of children and young people before and during the COVID-19 pandemic." *European child & adolescent psychiatry* 32, no. 6 (2023): 995-1013.

Madden, David. "Mental health in Ireland during the Covid pandemic: Evidence from two longitudinal surveys." *Plos one* 19, no. 3 (2024): e0279526.

Mandrekar, Jayawant N. "Receiver operating characteristic curve in diagnostic test assessment." *Journal of Thoracic Oncology* 5, no. 9 (2010): 1315-1316.

Melchior LA, Huba GJ, Brown VB, Reback CJ. A short depression index for women. *Educational and Psychological Measurement*. 1993 Dec;53(4):1117-25.

Mullainathan, Sendhil, and Jann Spiess. "Machine learning: an applied econometric approach." *Journal of Economic Perspectives* 31, no. 2 (2017): 87-106.

Nattino, Giovanni, Stanley Lemeshow, Gary Phillips, Stefano Finazzi, and Guido Bertolini. "Assessing the calibration of dichotomous outcome models with the calibration belt." *The Stata Journal* 17, no. 4 (2017): 1003-1014.

Orgad, Kfir. "The Impact of COVID-19 Lockdowns on Mental Health: A Comparative Study." *medRxiv* (2024): 2024-09.

Reiss, Franziska. "Socioeconomic inequalities and mental health problems in children and adolescents: a systematic review." *Social science & medicine* 90 (2013): 24-31.

Smyth E, Murray A. The Effect of Pandemic-Related Economic Disruption on Young Adolescents in Ireland. Children. 2022 Jul 12;9(7):1037.

Toffolutti V, Plach S, Maksimovic T, Piccitto G, Mascherini M, Mencarini L, Aassve A. The association between COVID-19 policy responses and mental well-being: Evidence from 28 European countries. *Social Science & Medicine*. 2022 May 1;301:114906.

Williams J, Greene S, Doyle E, Harris E, Layte R, McCoy S, McCrory C, Murray A, Nixon E, O'Dowd T, O'Moore M. Growing up in Ireland national longitudinal study of children. The lives of 9 year olds. The Stationery Office; 2009.

UCD CENTRE FOR ECONOMIC RESEARCH – RECENT WORKING PAPERS
SRAITH PÁIPÉAR OIBRE AN IONAID UM THAIGHDE EACNAMAÍOCHTA COBÁC

WP23/29 Judith M. Delaney, Paul J. Devereux:'Gender Differences in Teacher Judgement of Comparative Advantage' December 2023
WP23/30 Vincent Hogan, Patrick Massey: 'Different Strokes: Winning Strategies in Women's (and Men's) Big Bash Cricket.'
WP24/01 Ronald B. Davies, Lena S. Specht: 'Brexit and Foreign Students in Gravity' February 2024
WP24/02 Ronald B. Davies, Guohao Yang: 'A Comparison between Traditional and Knowledge Input Output Tables' February 2024
WP24/03 Tadgh Hegarty, Karl Whelan: 'Comparing Two Methods for Testing the Efficiency of Sports Betting Markets' February 2024
WP24/04 Matthew Amalitinga Abagna, Cecília Hornok, Alina Mulyukova: 'Place-based Policies and Household Wealth in Africa' February 2024
WP24/05 David Madden: 'The Trajectory of Obesity in a Cohort of Irish Children and their Mothers: An Application of Sequence Analysis' March 2024
WP24/06 Aline Bütikofer, Deidre Coy, Orla Doyle, Rita Ginja: 'The Consequences of Miscarriage on Parental Investments' March 2024
WP24/07 Håkan J. Holm, Margaret Samahita, Roel van Veldhuizen, Erik Wengström: 'Anchoring and Subjective Belief Distributions' April 2024
WP24/08 Judith M. Delaney, Paul J. Devereux: 'Gender Differences in Graduate Degree Choices' April 2024
WP24/09 Ciarán Mac Domhnaill: 'All hail? The impact of ride hailing platforms on the use of other transport modes' April 2024
WP24/10 Margaret Samahita: '"Luxury beliefs": Signaling through ideology?' June 2024
WP24/11 Alan de Bromhead,Seán Kenny: 'Irish Regional GDP since Independence' June 2024
WP24/12 Ronald B. Davies, James R. Markusen: 'Capital Ideas: Modelling and Measuring Factors in the Knowledge Capital Model' July 2024
WP24/13 Karl Whelan: 'Samuelson's Fallacy of Large Numbers With Decreasing Absolute Risk Aversion' July 2024
WP24/14 Cormac Ó Gráda: 'H1N1 and WW1: The Spanish Flu and the Great War' July 2024
WP24/15 Benjamin Elsner, Eoin T. Flaherty, Stefanie Haller: 'Brexit Had no Measurable Effect on Irish Exporters' August 2024
WP24/16 Eoin T. Flaherty: 'Are workers with multinational experience a determinant in startup success?' August 2024
WP24/17 Timothy G. Conley, Morgan Kelly: 'The Standard Errors of Persistence' October 2024
WP24/18 Zilong Li, Xi Chen, Zuzanna Studnicka: 'Have you eaten? The long-run impact of the Great Leap Famine on recent trade' November 2024
WP24/19 Karl Whelan: 'On Estimates of Insider Trading in Sports Betting' December 2024
WP25/20 Ciarán Mac Domhnaill: 'Driving over the hill: Car intensity during structural transformation' December 2024
WP25/01 Judith M. Delaney, Paul J. Devereux: 'Levelling the Playing Field? SES Differences in Graduate Degree Choices' February 2025
WP25/02 Zilong Li: 'International and Domestic Border Effects in China: Multilateral Resistances, Trade Substitution Patterns and Linguistic Differences' March 2025
WP25/03 Karl Whelan: 'The Gambler's Ruin with Asymmetric Payoffs' March 2025

UCD Centre for Economic Research
Email economics@ucd.ie