Feeling Useless: The Effect of Unemployment on Mental Health in the Great Recession

Lídia Farré
Francesco Fasani
Hannes Mueller

July 2015

Barcelona GSE Working Paper Series

Working Paper nº 838

Feeling Useless: The Effect of Unemployment on Mental Health in the Great Recession*

Lídia Farré
Universitat de Barcelona, IAE (CSIC) and MOVE

Francesco Fasani
Queen Mary University of London, IAE (CSIC) and MOVE

Hannes Mueller
IAE (CSIC), Barcelona GSE and MOVE

July 15, 2015

Abstract

This article documents a strong connection between unemployment and mental disorders using data from the Spanish Health Survey. We exploit the collapse of the construction sector to identify the causal effect of job loss. Our results suggest that an increase of the unemployment rate by 10 percent due to collapse of the sector raised mental disorders in the affected population by 3 percent. We argue that the large size of this effect responds to the fact that the construction sector was at the centre of the macroeconomic shock. As a result, workers exposed to the negative employment shock faced very low chances of re-entering employment. We show that this led to long unemployment spells, hopelessness and feelings of uselessness.

*We thank Ines Black for excellent research assistance and David Card for comments on an earlier draft. We acknowledge financial support from the RecerCaixa programme 2012. Mueller acknowledges financial support from the Ramon y Cajal programme and Spanish Ministry of Economy and Competitiveness, through the Severo Ochoa Programme for Centres of Excellence in R&D (SEV-2011-0075). All errors are ours.

1 Introduction

The Great Economic Recession which started with a financial crisis in 2007 had severe effects on the Spanish labor market. In particular, the unemployment rate followed a dramatic path, going from about 8 percent in 2007 to more than 25 percent in 2011. The construction sector was hit the hardest: more than 60 percent of all jobs in this sector were lost by 2013.[1]

This article shows that the unemployment suffered by the affected groups led to a drastic relative deterioration of their mental health. Figure 1 presents measures of mental well-being by employment status taken from the Spanish National Health Surveys of 2006 and 2011. Unemployed workers are clearly in worse health than their employed counterparts. They are less self-confident, appear overwhelmed by their problems and report markedly higher diagnosed mental disorders. However, these correlations come from cross-sectional evidence and are, therefore, uninformative about the underlying direction of causality. That is, mental disorders such as depression or chronic anxiety could be the result of unemployment, but it could also be that poor mental health leads to job loss or the inability to find new employment.

The Spanish economic recession offers a unique setting to study the causal relationship between unemployment and health. First, the deterioration of employment opportunities was directly linked to workers' exposure to the construction sector. Since the burst of the real estate bubble at the end of 2007, 3.8 million jobs have been lost: a third of them in construction. Second, the high concentration of job destruction in this sector, where workers with little education had been attracted by a decade of expansion, made unemployment a very hard trap to escape. Hence, the negative labor demand shock resulting from the collapse of the housing market resulted in exogenous job loss followed by a very low re-employment probability for the most affected workers. The nature of this economic episode therefore allows us to identify the effect of unemployment on health net of the biases resulting from the non-random selection of workers in and out of unemployment.

Our instrumental variable estimates suggest an important negative effect of unemployment on mental health, while non-robust findings appear on other health outcomes, including death rates. We also find that the IV estimates are much larger than those suggested by Figure 1 or the OLS regressions. In addition, the impact of unemployment on health is strongest for the last waves of the Health Survey data, i.e. in the years following the collapse of the construction sector.

[1] Author's calculations using the Spanish Labor Force Survey. See also Figure A1 in the appendix.

We argue that these findings respond to our identification strategy that relies on construction workers trapped into unemployment for a long time.

In the following section we review the related literature. Section 3 presents evidence on the changes in unemployment and unemployment duration with a focus on the construction sector. Section 4 discusses our data sources and section 5 provides a first look at the data. Section 6 introduces the empirical model and discusses our identification strategy. Section 7 presents our estimation results and some robustness checks. This is followed by some concluding remarks.

2 Related Literature and contribution

There is abundant evidence of a quantitatively large association between many economic indicators including income, wealth and employment status and a variety of health outcomes such as mortality, cardiovascular diseases or mental disorders (Ruhm 2000,2005). However, a heated debate remains about the direction of causality and about why the association arises. In this section we review the literature on the relationship between mental well-being and unemployment.

Psychologists and sociologists have long argued that unemployment damages mental health and a number of theories have been proposed to account for this relationship. For example, Jahoda (1982) and Warr (1987) argue that unemployment negatively affects mental health as it prevents a person from obtaining the non-monetary benefits of work such as a structured day, shared experience and opportunities of creativity and mental development. Alternatively, Erikson (1959) in his life-span development theory postulates that healthy emotional well-being among prime-age adults depends on the capacity to economically contribute to the family and, more generally, the society. In this sense, unemployment is harmful to mental health. Finally, those who blame themselves for undesirable happenings such as involuntary joblessness are likely to experience feeling of "helplessness" (Seligman, 1975) which damages mood and self-perception. Thus, for these persons, unemployment is expected to hamper mental health. Particularly relevant for our study is the phased response in emotional well-being found by Hill (1977) and others. In the first stage of the shock the individual is still optimistic. In the second stage, when efforts to obtain work fail, the individual becomes pessimistic and suffers active distress. In the third stage, the unemployed become fatalistic and adapts to the new state. Helplessness becomes acute.

A large body of literature reports a negative association between unemployment and a variety of health measures. At least three different paths can lead to the observation of a less healthy stock of unemployed compared to the employed. First, ill workers are more likely to become unemployed (García-Gomez et al. 2011). Second, there is evidence that poor health causes longer unemployment spells (Stewart, 2001). Finally, unemployment itself can lead to a deterioration of health. We focus on this third channel.

Some previous studies have employed panel data to estimate the effect of unemployment on health while controlling for unobserved time-invarying heterogeneity. However, this strategy cannot rule out the presence of health shocks that simultaneously affect health and employment status. Plant closures have been used as an alternative identification strategy to partially address the problem of reverse causality (Salm, 2009; Sullivan and von Wachter, 2009). Plant closure can well identify the short-term effects of unemployment as it represents an exogenous shock to the unemployment entry probability. However, the identification of the long-run effects will be tainted by the presence of selection effects into re-employment. For example, Stewart (2001) finds that individuals in poor health do have longer unemployment spells. Alternatively, it could also be that individuals who (expect to) suffer most from unemployment were more likely to try harder to escape it. We will argue that our identification strategy, based on the massive destruction of jobs in construction, will allow us to estimate the long-term effect of being jobless net of selection biases.

Our paper is also related to another stream of the literature that has examined the relationship between health and aggregate economic conditions, in particular unemployment. In a series of influential papers Ruhm (2000, 2003, 2005) finds that aggregate mortality is strongly procyclical, but that mental health (measured by the suicide rate) deteriorates during economic downturns. In the happiness literature, Clark and Oswald (1994) and Di Tella et al. (2001) document that higher levels of unemployment are linked with lower reported happiness. Using individual micro data, several studies also find that unemployment has a positive effect on suicide, depression, physician consultations, illness episodes, and substance abuse (see, among others, Dooley et al. 1996; Burgard et al. 2005).

3 The Spanish Economic Crisis

In this section we describe the main aspects of the economic crisis in Spain. We stress two main features. First, the negative shock to employment opportunities was mainly concentrated in the construction sector. Second, individuals who lost their jobs faced an extremely adverse labor market so that unemployment duration in the affected population increased dramatically.

Figure 2 shows the evolution of the average unemployment rate for provinces grouped according to the size of the construction sector in 2006 (i.e. large or small).[2] From the graph it is clear that the developments on the labor market between 2000 and 2012 were disastrous, as the unemployment rate dramatically skyrocketed over the period. Moreover, the shock was particularly severe in the group of regions with large levels of construction. Notice that until 2007 both groups were reducing unemployment almost in parallel. By 2010 unemployment in provinces with a large construction sector was almost 5 percentage points higher.

Figure 3 highlights the connection between the size of the construction sector and the incidence of unemployment. The y-axis shows the change in the unemployment rate between 2006 and 2011 in the 52 Spanish provinces. In the x-axis we show the share of employment in the construction sector over the total active population in 2006, before the crisis hit in 2007. The figure clearly shows that the largest increase in unemployment has been in those regions where employment in construction was the highest before the crisis. Some provinces had almost 1/5 of their active population employed in construction when the housing market collapsed. Five years later unemployment had risen by a similar amount. [3] In contrast, in regions with less construction, the unemployment rate suffered a much less pronounced increase.

In addition to the dramatic increase in the number of unemployed workers, the crisis in the Spanish labor market has also been characterized by an extremely low re-entry probability after job loss. Figure 4 reports the share of short term (less than 12 months) and long term unemployment over the active population in Spain. Until the Great Recession both rates were slightly decreasing. Long term unemployment in particular decreased from over 7 percent of the active population in 2000 to under 4 percent in 2007. In 2008, short-term unemployment increased drastically by about 2.5 percentage points. The following year the short-term rate increased again by 3.5 percentage points and stabilized thereafter. Long-term unemployment remained stable in 2008 but increased by about 2.7 percentage points in 2009 and by 2.1 percentage points in 2010. The long-term rate increased in all the following years and stood at 15 percent in 2013. This meant that the vast majority of individuals that lost their work in 2008 and 2009 did not find a job afterwards.

Individual reports on unemployment duration also reveal this change in the labor market.

[2] A province has a large construction sector if the share of employment in construction over total employment is above the mean value.

[3] For instance, in Tenerife the share of workers in the construction sector was 21% of the employed population in 2006. The unemployment rate increased from 8% to 30% in the province between 2007 and 2011.

Figure 5a shows the distribution of unemployment duration in the National Health Survey (Encuesta Nactional de Salud or ENS) sample for the years 2006 and 2011. Duration of unemployment changed dramatically between these two years. In 2006 about half the unemployed workers experienced spells that lasted less than 6 months. As a result of the economic downturn this group increased slightly from over 5 percent of the active population in 2006 to about 7 percent in 2011. Most of the additional unemployed, however, experienced longer spells. In particular, the group with unemployment spells of more than two years more than tripled in size from about 2 percent in 2006 to almost 8 percent in 2011.

Construction workers were most affected by long-term unemployment. The ENS asked individuals in both 2006 and 2011 whether their current or last employment was in construction. Figure 5b shows unemployment duration in this group, again as percent of the active population. Unemployment in this group increased particularly strongly and an overwhelming majority of the additional unemployed was without employment for longer than a year in 2011. While the Spanish Labor Force Survey does not provide data on long-term unemployment by sector it should be clear from these numbers that, if anything, the pattern displayed in Figure 4 should be even more extreme for construction workers.

4 Data

This paper employs data from two main sources. Information at the individual level on health and employment status as well as other socioeconomic characteristics is obtained from the different waves of the National Health Survey (ENS). This survey exists for different years between 1987 and 2011. In the years 1987, 1993, 1995, 1997 and 2001 the survey was conducted by the Centro de Investigaciones Sociologicas (CIS), an independent entity assigned to the Spanish Ministry of the Presidency. In 2003, 2006 and 2011 the Ministry of Health was in charge of it. While the different waves of the survey are designed to analyze the health status and practices of the Spanish population, the questions are, in general, not comparable across time. Most part of our analysis focuses on the comparison between the year 2006 (just before the economic collapse) and 2011 (in the middle of the economic downturn), using the last two waves of the survey. The questions in these two surveys are almost identical. We present additional results for a general health measure and our main measure of mental disorder which are both available for the waves

2001, 2003, 2006 and 2011.[4]

We exclude individuals that are under 17 or older than 64. Unless stated otherwise we only look at the active population, which implies that we exclude students, disabled and pensioners. Table 1 presents the descriptive statistics for the variables we use from the NHS. The sample size is about 46,000 in the larger sample and 25,000 in the last two waves 2006 and 2011 but varies slightly depending on the specific health question we consider.

The ENS survey provides very detailed questions on aspects of health. First, respondents are required to provide a self assessment of their general health status, classifying it in very good/good/bad/very bad health. We recode this variable giving values of 1 to reports of very good or good health. Second, respondents are asked whether they received a diagnosis from a doctor for a set of different illnesses (e.g. chronic back pain; chronic headache; heart attack, stroke, etc.). Of particular interest for us is the question regarding whether the respondent has been diagnosed with a mental disorder (i.e. depression or chronic anxiety). Third, a measure of self-reported mental health is obtained by asking respondents whether they suffer from some mental disorder. Table 1 shows that the sample average for general good health and mental disorder are fairly stable across samples, 80 percent of the active population report good health. Mental disorders are reported by about 8 percent on average. Health was improving slightly between 2006 and 2011. The percentage of individuals reporting good health increased from 76 percent to 81 percent, for example. Reported mental disorders fell from 9 percent in 2006 to 7 percent in 2011.[5,6] However, this positive trend was not uniform. Individuals associated with the construction sector reported slightly worse mental health on average in 2011 than in 2006 (see Appendix Table A1).

[4]The question regarding general health status is the same in 2001, 2003, 2006 and 2011: *"Over the last 12 months, would you say your health has been? Very good, Good, Average, Poor or Very poor"*. In 2006 and 2011 the same question is asked regarding diagnosed mental disorders: *Have you ever been diagnosed by your doctor with chronic depression, anxiety or any other mental disorder?*. In 2001 and 2003 the question is: *Are you currently diagnosed by your doctor with chronic depression, anxiety or any other mental disorder?"*

[5]This is in line with data on death rates. Death rates from the four main sicknesses (cancer, respiratory diseases, infectious diseases and cardiovascular diseases) were falling throughout the 2000s including the crisis years - see Appendix Figure A2.

[6]Data on death rates were taken from the population census between 2006 and 2011. Death episodes were presented by cause of death, under the following code: 1-Cancer, 2-Respiratory Disease, 3-Infectious Disease, 4-Cardiovascular Disease, 5-Traffic Accidents, 6-Other Accidents, 7-Suicide, 8-Homicide.

The ENS waves of 2006 and 2011 conduct a special survey of twelve questions related to mental well-being. The questions are part of the General Health Questionnaire (GHQ) which was developed as a screening instrument for psychiatric illness. Responses in this survey are coded between 0 and 3, where 3 is always the worst outcome, 1 is the default and 0 indicates a better than usual state of mind. To make interpretation easier we recoded the variables with 0 or 1, where 1 indicates a response worse than usual. The questions can be grouped in three categories. The first category are stress-related indicators and questions for general well-being (for example, In the last couple of weeks have you: lost much sleep over worry?; felt constantly under strain?; been feeling reasonably happy, all things considered?). The second category proves the decision-making capacity of individuals (for example, In the last couple of weeks have you: been able to concentrate on whatever you are doing? felt capable of making decisions about things?). The third category contains questions about the individuals self-perception (for example, In the last couple of weeks have you: felt you were playing useful part in things?; being thinking of yourself as a worthless person?). While it could be argued that some of these answers simply capture general well-being it is harder to claim the same for other questions (for example, "Have you been able to concentrate on whatever you are doing?"). Between 2006 and 2011 the mean of these measures remained constant at a around 10 percent.

The second data set employed in our estimation is the Spanish Labor Force Survey (Encuesta de Población Activa - EPA). This survey is an ongoing research carried out every quarter and it targets households. Its main objective is to obtain data on the labor force and the various categories (employed and unemployed persons), as well as the population out of the labor market (inactive persons). The initial sample includes 65,000 interviewed households per quarter, which implies approximately 180,000 people.

Additionally, we used population data to build rates (death rate, unemployment rate). Population data was gathered from the National Statistics Institute (Instituto Nacional de Estadistica or INE), specifically from the municipal register (continuous statistics) from 2006 to 2011. We used a disaggregation by province.

5 Health and Unemployment: Descriptive Evidence

We start our empirical analysis using information from the National Health Survey to investigate the correlation between unemployment and several health indicators. We estimate the following OLS regression:

$$health_{ipt} = \alpha u_{ipt} + \boldsymbol{\beta} X_{ipt} + \theta_p + \eta_t + \nu_{ipt} \tag{1}$$

where the dependent variable, $health_{ipt}$, is a measure of health for individual i, residing in region p at time t. The model includes a dummy variable to capture whether the respondent is unemployed, u_{ipt} (our main regressor of interest), a vector of individual socioeconomic characteristics, X_{ipt}, fixed effects at the region, θ_p, and year level, η_t, and an error term ν_{ipt}.

Table 2 shows the first set of results. In the first column, the dependent variable is an indicator that takes value 1 if the respondent declares to be in *good* or *very good* health and 0 otherwise. This question is common to all the waves of the survey, and thus we include in estimation all the observations since 2001.[7] A gender dummy and an indicator for being younger than 40 are included as additional controls. All dependent variables are divided by their standard deviation to provide some comparability across survey questions. The estimated coefficient on the unemployment dummy reported in column (1) implies that the unemployed have 20 percent of a standard deviation worse health than the employed. There is also strong evidence that men and young report better health. These results still hold when controlling for education categories or finer age groups.

The remaining columns of Table 2 display the results of an alternative empirical specification employed in most of the paper. This new specification is based on a cell-level panel where cells are defined by three variables: age, sex and province of residence. The idea now is to compare changes in health outcomes across time holding a combination of individual characteristics (i.e. age and gender) and geography fixed. Accordingly, we include cell fixed effects θ_c in the specification in equation (1) and estimate the following model:

$$health_{ipt} = \alpha u_{ipt} + \theta_c + \eta_t + \nu_{ipt}. \tag{2}$$

The amount of data we have does not allow for a very fine-grained distinction in age groups across provinces. Thus in our main specification we distinguish individuals that are older or younger than 40. Cells, c, are therefore defined by:

$$c : \{under40, province, male\}$$

which gives us $2 \times 51 \times 2 = 204$ cells.

The new specification is quite demanding as it now allows average health levels to vary across provinces for combinations of age and sex. Column (2) shows that our more stringent specification

[7] The same results hold if we include earlier waves.

provides the same results regarding general health. Using the cell specification in equation (2), column (3) shows that mental disorders are 16 percent of a standard deviation more likely among the unemployed. Column (4) restricts the sample to the most comparable survey waves in 2006 and 2011 and results barely change.[8] Other illnesses like chronic headaches and heart attacks are also more likely among the unemployed. However, here the magnitudes are much smaller. Heart attacks, for example, increase by about 6 percent of a standard deviation with unemployment.

The estimates in Table 2 highlight a clear correlation between mental health and, to a lesser extent, health in general and unemployment; however they are uninformative about which direction causality runs. Indeed, the OLS estimates of unemployment status on mental health mix two aspects. On the one hand, those who are in unemployment may have a different level of mental health than those who are employed. This will be the case if pre-existing mental health problems correlate with a higher likelihood of being fired and/or if mental disorders make job search harder.[9] On the other hand, entering (or remaining) unemployed may lead to isolation and economic stress, which can then trigger or amplify mental disorders. Only this latter effect is the causal impact of unemployment on health; the parameter we are after in our estimation. To this end, we employ an instrumental variable strategy based on the massive destruction of jobs in the construction resulting from the bursting of the Spanish housing bubble.

6 Empirical Strategy

6.1 Theoretical Discussion

Our empirical analysis exploits the features of the recent Spanish economic to identifying the causal effect of unemployment on health. We employ a two-stage least square estimation technique where the unemployment variable in equation (2) is instrumented using an individual's exposure to the collapse of employment opportunities in the construction sector. We argue that this instrument, in the context of the Spanish recession, satisfies two separate and important conditions: i) job losses are exogenous to unobserved individual characteristics; ii) re-entry into employment is almost impossible. We now discuss these two assumptions theoretically in a static

[8] To get a feel for the magnitude it is useful to note that the base mean of mental disorders is 8 percent while 16 percent of a standard deviation are about 5 percent.

[9] This latter pattern could be driven either by screening of employers or by a reduced capacity of effectively looking for jobs among the mentally ill.

framework. For a derivation in a dynamic framework see the appendix.

To analyze the relevance of these two conditions let us first assume a situation where the effect of unemployment on health is homogeneous in the population. Accordingly we can estimate the following equation:

$$h_{it} = \alpha_0 + \alpha u_{it} + \mu_i + \epsilon_{it} \tag{3}$$

where h_{it} is (mental) health status of individual i at time t, u_{it} is a dummy equal one if the individual i is unemployed at time t, μ_i is an individual fixed effect and ϵ_{it} is an error term.[10] If being unemployed negatively affects an individual's health, we should expect the coefficient on the unemployment indicator to be negative ($\alpha < 0$).

One could assume that $cov(u_{it}, \mu_i) = 0$, but there are reasons to expect $cov(u_{it}, \mu_i) \neq 0$. In particular, under the realistic assumption that healthier individuals are less likely to be unemployed (i.e. if productivity is increasing in health, employers prefer hiring healthier individuals), then $cov(u_{it}, \mu_i) < 0$. Under this assumption, the OLS estimator can be written as:

$$\begin{aligned} \alpha_{OLS} &= E(h_{it}|u_{it} = 1) - E(h_{it}|u_{it} = 0) \\ &= \alpha + E(\mu_i|u_{it} = 1) - E(\mu_i|u_{it} = 0). \end{aligned} \tag{4}$$

If individuals who are unemployed have on average lower health than those employed, we will have that $E(\mu_i|u_{it} = 1) - E(\mu_i|u_{it} = 0) < 0$. Under this assumption the OLS estimator is expected to be larger in magnitude than an IV estimator that manages to retrieve the actual parameter α (i.e. $|\alpha_{OLS}| > |\alpha|$).

However, there is no reason to expect the effect of job loss to be homogenous in the population. On the contrary, we can expect different people to react differently to the experience of being unemployed. Being laid off can be a psychologically devastating experience for some, whereas for others it may just represent an unfortunate accident in life. Clearly, several factors will determine the impact that being unemployed has on each individual. Workers who can rely on savings, family wealth or spouse's income, for instance, will not have to immediately worry about the economic consequences that losing a job may have. Beyond short-term concerns generated by the loss of income, the magnitude of the mental impact of unemployment will also depend on individual psychological traits such as self-esteem and self-confidence, on whether the

[10] For simplicity we remove the geographic dimension out of estimation in this part of the discussion.

individual experienced unemployment before, on the social stigma that the individual attaches to the unemployment status, etc. Further, expectations should play a crucial role. The fact of being laid off will generate more stress the less expected the event was, and stress should increase if the individual deems it difficult to find a new job in the near future.

Under the assumption that the effect of unemploment is heterogeneous in the population, equation (3) can be written as:

$$h_{it} = \alpha_0 + \alpha_i u_{it} + \mu_i + \epsilon_{it} \tag{5}$$

where now the coefficient α_i varies at the individual level. Using the notation in the policy evaluation literature, we have:

$$\begin{aligned} h_{it} &= \alpha_0 + \alpha^{ATE} u_{it} + (\alpha_i - \alpha^{ATE}) u_{it} + \mu_i + \epsilon_{it} \\ &= \alpha^{ATE} u_{it} + e_{it} \end{aligned} \tag{6}$$

where ATE refers to the *average treatment effect* of unemployment in the population. That is, $\alpha^{ATE} = E(\alpha_i)$ and $e_{it} = (\alpha_i - \alpha^{ATE}) u_{it} + \mu_i + \epsilon_{it}$.

Following with this notation, the expected health among the employed can be expressed as:

$$E[h_{it}|u_i = 1] = \alpha_0 + \alpha^{ATE} + E[(\alpha_i - \alpha^{ATE})|u_i = 1] + E[\mu_i|u_i = 1], \tag{7}$$

and among the unemployed:

$$E[h_{it}|u_i = 0] = \alpha_0 + E[\mu_i|u_i = 0]. \tag{8}$$

Therefore, the OLS estimator is:

$$\begin{aligned} \alpha_{OLS} &= E[h_{it}|u_i = 1] - E[h_{it}|u_i = 0] \\ &= \alpha^{ATE} + E[(\alpha_i - \alpha^{ATE})|u_i = 1] + E[\mu_i|u_i = 1] - E[\mu_i|u_i = 0]. \end{aligned} \tag{9}$$

The selection bias in the presence of heterogeneity has two components. First, the term $[E[\mu_i|u_i = 1] - E[\mu_i|u_i = 0]]$ which also appears in the case of homogeneous effects and can be expected to be negative if healthier individuals are less likely to be unemployed. Second, the term $E[(\alpha_i - \alpha^{ATE})|u_i = 1]$ which reflects the possibility of differences across individuals in the effect of unemployment. In this context, we can expect individuals who would suffer the most from unemployment to be less likely to be unemployed. Indeed, these are the individuals who have stronger incentives to exert maximum effort to keep their job (or to find one, if they are unemployed) and to lower their reservation wage to avoid unemployment. Individuals with higher potential (mental) health loss from unemployment, therefore, will have a lower probability of entering unemployment if employed and higher probability of exiting unemployment if unemployed. This implies that we should expect unemployed individuals to have α_i above the average in the population, that is: $E[(\alpha_i - \alpha^{ATE})|u_i = 1] > 0$.[11]

Therefore, in the presence of heterogeneity we have two sources of bias in the OLS estimator and they have opposite signs. Differently from the homogenous case, it is now unclear whether the OLS estimator over- rather than under-estimateS the causal parameter of interest. The bias will depend on whether selection into and out of unemployment correlates with health status or the health loss in unemployment.

In order to retrieve the causal effect of unemployment one would need an instrument that is uncorrelated with both the unobservable health status of workers, μ_i, and with the unobservable individual "health effect" from being unemployed, α_i. In other words, one would need an exogenous shock that pushes individuals into unemployment irrespective of their unobservables characteristics. The literature has proposed to use plant closures as instrument in this context (see, for example, Salm (2009)). Indeed, when a plant shuts down all employees are generally laid off. For these workers, the entry into unemployment is completely orthogonal to their unobservable individual characteristics. However, if the initial sample of laid-off workers is (arguably) as good as random, nothing prevents these workers to actively search for new jobs. The problem appears then when the effect of unemployment on health is estimated at different points in time after the plant is closed. It may well be that those who are still unemployed suffered less from this job market status. Accordingly, the plant closure instrument removes the selection on unobservable health status (i.e. $cov(u_{it}, \mu_i) = 0$) but fails to remove the bias due to selection on the idiosyncratic effect of unemployment on health (i.e. $cov(u_{it}, \alpha_i) \neq 0$). Given that we expect this latter component to be positive, IV estimates obtained with the plant closure approach may underestimate the magnitude of the causal effect of unemployment on health.

The instrument we proposed in this paper based on the collapse of the construction sector follows a logic similar to the plant closure instrument but has one important differences. The bust of the housing bubble in Spain after the 2007 economic recession represented the shut down of almost an entire sector. Thus similarly to the plant closure case we have an exogenous push into unemployment of individuals irrespective of their unobservable characteristics. However, we

[11] Recall that we expect α^{ATE} to be negative

now have a situation where exiting unemployment is very hard, if not impossible. Indeed, while workers negatively affected by an idiosyncratic shock (i.e. a plant closure) can quickly find a new job in any other firm, workers laid off by the shut down of an entire sector will find themselves trapped in unemployment unless they manage to change sector. As documented in section 3, the collapse of the construction sector led to both a large increase in unemployment and to a dramatic increase in its duration, with exit rates from unemployment being driven close to zero. The lack of unemployment exit possibilities removes - or, at least, greatly reduces - the concern that endogenous sorting out from unemployment prevents us from identifying the causal effect of interest.

6.2 Construction of the Instrument

The previous discussion highlights that we need a variable that captures exogenous job loss and homogenous re-employment probabilities across workers. The evolution of employment in the Spanish construction sector can serve as an instrument for both. First, the collapse of the sector meant that individual fortunes were driven by an exogenous shock. Between 2007 and 2012, employment fell by more than 60 percent.[12] Many businesses had to close: bankruptcies in construction shot up from just around 200 per year in the period 2005 to 2007 to around 1500 per year in the period 2008 to 2010, and they reached 1900 in 2011. The increase in bankruptcies was not only in absolute terms but also in relative terms: about 33 percent of all bankruptcies in Spain between 2008 and 2010 were by companies in construction.[13] This suggests that if we use employment in construction as an instrument for unemployment we will be capturing job losses due to plant closure from the year 2007 onwards. Workers in construction were often unskilled with a training very specific to the sector. Thus as a result of the collapse they had a hard time in finding a new job and were trapped in unemployment for a long time.[14]

To form the instrument we employ the exposure of different groups to the construction sector. The idea behind our identification strategy is to use changes in demand for labor at the aggregate level as an instrument for unemployment at the cell level. As before, we use cells spanned by three characteristics: age, sex and province of residence:

$$c = \{under40, province, male\}.$$

[12] See appendix figure A1.

[13] Source: Spanish Statistical Office.

[14] See the evolution of employment duration in the construction sector over the period in Figure 5b.

For these cells we construct employment shares by 9 industries, j, in 2000. We refer to these shares as $s_{c,j,2000}$. As a second step we calculate the aggregate employment growth in industry j in year t at the national level as:

$$g_{jt} = (E_{j,t} - E_{j,t-1}) / E_{j,t}.$$

We focus on employment growth as this gives us variables without a time trend. Both our first stage and second stage results are robust to using employment levels.

Figure 6 shows employment growth in Spain. We plot the average employment growth for all sectors with a dashed line and employment growth in construction with a solid one. The picture shows that until 2007 employment was growing in Spain, but the boom was particularly large in construction where growth was above average in all years. However, in 2007 the shock hit and employment fell across the board. The shock was particularly strong in construction where employment shrank by more than 20 percent in 2009 and growth was below -10 percent in all years after 2007. As a result, more than 60 percent of all jobs that existed in construction in 2007 were lost in the following years. This was a very drastic development even when compared to the generally dramatic change in the Spanish labor market, where most sectors shed about 15 percent of employment after 2007.

We use the interaction between the share of employment in construction at the cell level in 2000 (i.e. $s_{c,constr,2000}$) and the annual employment growth in construction at the national level (i.e. $g_{constr,t}$) as our main instrument for cell unemployment. That is:

$$construcIV_{c,t} = s_{c,constr,2000} * g_{constr,t}.$$

As an alternative we also employ an instrument based on total employment growth:

$$employmentIV_{c,t} = \sum_j s_{c,j,2000} * g_{j,t}.$$

The idea behind this instrumental variable approach is that aggregate changes in employment are not driven by cell specific characteristics. Moreover, its interaction with the industry composition in 2000 ensures that the exposure of cells to construction is pre-determined.

Our first stage regression then follows

$$u_{it} = \delta \times construcIV_{c,t} + \theta_c + \eta_t + \nu_{it} \tag{10}$$

where the unemployment status of individual i, in cell c, at time t, is regressed on the cell-specific instrument. The regression includes a full set of cell fixed effects, θ_c, and year fixed effects, η_t. In this specification the parameter δ captures the change in unemployment for individuals which can be explained by the change in job opportunities in construction.

7 Results

7.1 First-Stage Estimates

Table 3 reports variations of the first stage regression in equation (10). Column (1) to (6) report results for all the waves in the NHS 2001, 2003, 2006 and 2011. Column (1) employs all sectors in forming the instrument. There is a clear negative correlation between employment growth and the level of unemployment at the cell level. In column (2) the predicted level of total employment growth is divided into the construction and all other sectors. The results show that employment in construction is a much better predictor of unemployment. This is consistent with the much more rapid decline in the employment of this sector observed in Figure 6. The estimates in column (3) only include employment growth in construction and suggest that the shrinking of the 15 percent of employment prior to 2011 led to an unemployment rate of more than 30 percent in cells that had all their employment in the construction sector in 2000. This finding is robust to various modifications with respect to the definition of cells (columns (4) and (5)) and also does not change if we focus on just the last two waves in 2006 and 2011 (column (6)). However, in column (7) we show that the instrument is relatively weak if we just study the first three waves 2001, 2003 and 2006.

In our main analysis we employ as a first-stage the results in column (6). It provides fitted values of the unemployment rate of up to 58 percent. The average change across the two waves is an increase of 12 percentage points and the maximum increase is 24 percentage points.[15] The group with the biggest increase are men below 40 in provinces with large construction sectors.

7.2 Main Results

Table 4 presents the main results. The estimates are obtained from the second-stage regression:

$$health_{ict} = \alpha \hat{u}_{ct} + \theta_c + \eta_t + \nu_{ict} \tag{11}$$

where $\hat{u}_{ct}$ is the predicted unemployment from equation (10). In this regression we need to cluster at the cell level since all variation in $\hat{u}_{ct}$ comes from the cell level.[16] As before we control for cell and year fixed effects.

[15] In appendix Figure A3 we report kernel densities of the fitted values in column (6).

[16] This is despite the fact that we instrument for unemployment at the individual level. We return to this point below.

The results in Table 4 are obtained from the comparison of health in the two latest waves, 2006 and 2011. We reweigh all the dependent variables according to their standard deviation.[17] We find a strong and negative effect on reports of general good health (-0.74 standard deviations) and large effects on mental disorders with and without diagnosis by a doctor (about 1.1 standard deviations). The estimates indicate that a 10 percentage point increase in unemployment driven by the exogenous shock, increases mental disorder by about 3 percentage points.

The remaining columns in Table 4 confirm the findings on mental health using the GHQ questionnaire on mental disorders. Remember that all questions here are coded such that positive coefficients indicate a "worse than usual" answer. Column (4) reports that unemployment leads to an increase of 0.9 standard deviations in the mean score across all categories in the questionnaire. On each question we find a positive and fairly large coefficient. However, only a few are significantly different from zero. In particular, the unemployed are 1.3 standard deviations more likely to report that they feel constantly under strain and 0.9 standard deviations more likely to report that they do not feel a useful part of society. There is also some evidence that the unemployed are more likely to feel that they cannot overcome their difficulties and are unable to concentrate.

Notice that the IV estimates of the effect of unemployment on mental health are much larger than the OLS reported in Table 3. This large size responds to the sub-sample of the population from where the effect is identified, namely construction workers. As discussed, employment in construction fell by about 60 percent between 2007 and 2013 and the large majority of those who lost their job in construction, 2.7 percent of the active population, slipped into unemployment spells that lasted longer than one year. Accordingly, while workers negatively affected by an idiosyncratic shock can quickly find a new job in any other firm, workers laid off by the shut down of an entire sector find themselves trapped into unemployment. Failure to re-enter employment for those who try hardest might have very high costs on mental health. This view is corroborated by the finding that affected individuals felt under strain and useless.

In light of our theoretical discussion the coefficient we identify with our instrument is a *Local Average Treatment Effect* (LATE). This effect is defined on a specific population of *compliers*: those workers who entered unemployment as a consequence of the collapse of the construction

[17] This means all coefficients can be interpreted as the impact of unemployment on the dependent variable in terms of its standard deviation.

sector. Note that the population of unemployed workers after the crisis hit Spain can be distinguished in two groups. A first group of unemployed workers called *always-takers:* they would have been unemployed even in the absence of the crisis. We can think of this sub-population as the workers who would have been unemployed even in normal times and we can expect them to be those who suffer relatively less from unemployment (i.e. those who have relatively low α_i). The second group of unemployed workers are those who were pushed into unemployment by the crisis, the *compliers*: these are individuals who would have been employed had the crisis not hit. We can therefore expect these individuals to have average α_i well above those of the *always-takers.* Given the characteristics of our IV strategy, we identify the average treatment effect precisely among this latest group of the population.

There is an alternative and complementary explanation to the large size of our IV results. In our previous discussion, we assume that the identifying condition of the instrument holds at the individual level. However, the variation in the instrument is only at the cell level. It could well be that the effects of high unemployment in a cell spilled over to the employed. This is reasonable as cells (i.e. male, under 40, province of residence) are precise enough to capture local labor markets. The treatment of unemployment is then literally at the cell not at the individual level. This interpretation would not violate the restriction assumption on the IV if the spill-over works through past experience of unemployment and the fear of (long term) unemployment amongst those who have work.[18]

7.3 Additional Results

Tables 5 and 6 report additional results. In Table 5 we report IV estimates of the effects of unemployment on other health outcomes. We find some weak evidence that chronic headaches become more likely as a result of becoming unemployed; but otherwise we find very few consistent results. This is interesting as it suggests that unemployment caused by the shock did not, yet, lead to a general deterioration of health. For example, the fact that the OLS results in Table 2 regarding stroke go away suggests that these were probably driven by reverse causality. In column (5) of Table 5 we show that the unemployed are more likely to use medicine. This is in line with the finding that general health and in particular mental health deteriorates.

[18] This interpretation is also consistent with the findings in Sullivan and von Wachter (2009); the employment shock of a plant closure affects individuals who find work later.

Finally, we analyze the effect of unemployment on suicides. Figure 7 reports the level of suicides per 100,000 population which we calculate from deaths and population numbers. Suicide rates were falling from 7.6 in 2000 to 6.6 (per 100,000) in 2011. However, the fall is not uniform but interrupted by two large waves. The second wave starts exactly in 2007. In Table 6 we confirm that the increase in suicides during this second period took place in those cells that were hardest hit by unemployment. To do this we take unemployment rates at the cell level and run a IV regression of ln(suicides) on unemployment which follows equation (11). The only difference to our main results is that we use unemployment rates from the EPA and therefore have yearly data for the period 2001-2011. Column (1) indicates that, overall, there is no consistent relationship between unemployment and suicides in the period 2001-2011. The positive association between unemployment and suicides only becomes apparent if we focus on the years after 2006. The relationship is then robust to the inclusion of ln(population) on the right hand side of equation (11), province time trends and modifications in the definition of cells. This result would suggest that an increase of the unemployment rate by 10 percentage points leads to an increase in suicides by about 45 percent. This is an increase of about 3 deaths in 100,000 population per year. However, this interpretation is problematic given the earlier peak which fell into a period of falling unemployment.

7.4 Robustness

We now present a number of robustness checks to our main results in Tables 7, 8 and 9. First, we use a different division in cells as introduced in Table 3. Our first alternative uses a finer distinction by age. That is:

$$c = \{under30, over50, province, male\}$$

and run the same regressions as in Table 4. Note that we now have 306 cells and control for many more cell fixed effects. Results are unaffected by this (Table 7). If anything the results from the GHQ survey strengthen. We then add college education as an additional dimension:

$$c = \{under40, province, male, college\}$$

and, again, control for cell fixed effects at this level (408 cells). Under this alternative definition the results are also not significantly affected (Table 8). The coefficient on the general health indicator drops and becomes insignificant while several variables in the GHQ are now estimated with more precision.

Table 9 presents a number of additional robustness checks just with our main measure of mental disorder diagnosed by a doctor. Columns (1) to (3) use variants of the IV variable. Column (1) uses total employment as an instrument. The estimated effect on unemployment slightly increases but we cannot reject that this coefficient is different from the one estimated in the main Table 3. Column (2) uses employment growth in the previous three years to instrument for unemployment. Results remain unchanged. This is also true if we just use employment levels or employment changes. Column (3) uses only variation at the province level, clustering also only at this level. We still find a positive coefficient but the standard errors are now much larger, and the coefficient becomes insignificant. Column (4) uses the unemployment rate at the cell level constructed from the Spanish Labor Force Survey instrumented using the predicted growth in employment. Our results are robust to this different way of looking at the data. In column (5) we add the inactive population (pensioners, students, individuals working from home) and our results on unemployment do not change.

In columns (6) to (8) we add the earlier waves in the NHS. In column (6) we estimate our preferred specification by adding the waves 2001 and 2003.[19] The coefficient drops slightly and is now only significant at 10 percent. This is in line with the idea that what drives our results is the extreme shock to employment opportunities between 2006 and 2011. The main benefit of adding more waves is that we can control for long term trends in health. In column (7) we control for province specific time trends and results remain the same. In column (8) we include in specification a time trend for men. This is based on the idea that our construction sector instrument could be capturing the relative movement of mental health between men and women. Our results strengthen under this alternative specification, suggesting that the instrument does not merely capture long term gender trends.

8 Conclusion

In this article we analyze the relationship between unemployment and mental health in the context of the severe economic crisis in Spain. We exploit the extreme circumstances in the labor market of construction workers to identify the causal effect of unemployment on health. We argue that job destruction as a result of the burst of the housing bubble represented an

[19] Remember that the questions regarding mental disorder were not exactly the same between 2001/2003 and 2006/2011.

exogenous shock to labor demand that affected both the probability of being laid off as well as that of re-employment. Accordingly, our instrumental variable approach is able to estimate the causal effect of unemployment on health net of workers' selection in and out of unemployment. Our IV estimates suggest that mental disorders in this group are almost 30 percentage points more likely than in the employed population. The large magnitude of this effect responds to the fact that identification comes from a group of workers that were unable to escape unemployment after the collapse of the construction sector.

Our findings raise the concern that a significant share of the Spanish labor force could get trapped in a cycle of skill mismatch and mental disorder. Long-term unemployment stood at 12 percent of the active population in 2012. The finding that this group is not only suffering from an income loss but from a loss of (mental) health is worrying on its own right. In addition, the combination of skill mismatch and the inability to search and embrace new labor market opportunities in such a large part of the population is a liability for the Spanish economy as a whole.

References

[1] Burgard, S.A.; J.E. Brand and J.S. House. (2005). Causation and Selection in the Relationship of Job Loss to Health in the United States.

[2] Card; David (2001) Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts Immigration. Journal of Labor Economics, Vol. 19, No. 1 (January 2001).

[3] Clark, Andrew (2003) "Unemployment as a Social Norm: Psychological Evidence from Panel Data," Journal of Labor Economics, University of Chicago Press, vol. 21(2), pages 289-322, April

[4] Clark, A. and A.J. Oswald. (1994). Unhappiness and unemployment. The Economic Journal 104 (424), pp 648-659

[5] Clark, Andrew & Knabe, Andreas & Rätzel, Steffen, 2010. "Boon or bane? Others' unemployment, well-being and job insecurity," Labour Economics, Elsevier, vol. 17(1), pages 52-61, January.

[6] Diette, T.M., A. Goldsmith, D. Hamilton and W. Darity. 2012. Causality in the Relationship between Mental Health and Unemployment. Upjohn Institute Press Book Chapter.

[7] Di Tella, R., R.J. MacCulloch, A.J. Oswald. (2001) Preferences over inflation and unemployment: evidence from surveys of happiness. American Economic Review, 91 (1), pp. 335-341.

[8] Dooley, D.; J. Fielding; and L.Levi. (1996). Health and unemployment. Annual Review of Public Health, 17, pp 449465.

[9] Douglas L. Miller & Marianne E. Page & Ann Huff Stevens & Mateusz Filipski, 2009. "Why Are Recessions Good for Your Health?," American Economic Review, American Economic Association, vol. 99(2), pages 122-27, May

[10] Ensminger, Margaret E., and David D. Celentano. 1990. "Gender Differences in the Effect of Unemployment on Psychological Distress." Social Science and Medicine 30(4): 469–477.

[11] Erikson, Erik H. 1959. "Identity and the Life Cycle." Psychological Issues 1(1): 50–100.

[12] García-Gómez, P., A.M. Jones and N. Rice. (2010). Health effects on labour market exits and entries. Labour Economics, 17, p.p. 62-76.

[13] Harrison, Richard. 1976. "The Demoralising Experience of Prolonged Unemployment." Department of Employment Gazette 84(4): 339–348.

[14] Hilary Hoynes & Douglas L. Miller & Jessamyn Schaller, 2012. "Who Suffers during Recessions?," Journal of Economic Perspectives, American Economic Association, vol. 26(3), pages 27-48, Summer.

[15] Izquierdo, M., J.F. Jimeno and A. Lacuesta (2013). "Spain: From (massive) immigration to (vast) emigration?", Banco de España Working Paper.

[16] OECD (2013) "Crisis squeezes income and puts pressure on inequality and poverty".

[17] Jackson, Paul R., and Peter B. Warr. 1984. "Unemployment and Ill-Health: The Moderating Role of Duration and Age." Psychological Medicine 14: 605–614.

[18] Jahoda, Marie. 1982. Employment and Unemployment: A Social-Psychological Analysis. New York: Cambridge University Press.

[19] Jahoda, Marie, Paul F. Lazarsfeld, and Hans Zeisel. 1933. Marienthal: The Sociography of an Unemployed Community (English translation, 1971). Chicago, IL: Aldine.

[20] Pritchett, Lant, and Lawrence H. Summers (1996) "Wealthier is Healthier," Journal of Human Resources, XXXI , 841–868.

[21] Ruhm, Christopher J. 2000. "Are Recessions Good for Your Health?" Quarterly Journal of Economics, 115(2): 617-50.

[22] Ruhm, Christopher J. 2003. "Good Times Make You Sick." Journal of Health Economics, 22(4): 637-58.

[23] Ruhm, Christopher J. 2005. "Mortality Increases During Economic Upturns." International Journal of Epidemiology, 34(6): 1206-11.

[24] Salm, M. 2009. Does job loss cause ill health? Health Economics, 18(9), pp. 1075-1089.

[25] Seligman, Martin, E. P. 1975. Helplessness: On Depression, Development and Death. San Francisco, CA: W. H. Freeman

[26] Stewart, J.M. (2001). The impact of health status on the duration of unemployment spells and the implications for studies of the impact of unemployment on health status. Journal of Health Economics, 20(5), pp. 781-796.

[27] Sullivan, Daniel G., and Till von Watchter. 2007. "Mortality, Mass-Layoffs, and Career Out comes: An Analysis using Administrative Data." National Bureau of Economic Research Working Paper 13626

[28] Warr, Peter B. 1987. Work, Unemployment and Mental Health. Oxford, United Kingdom: Oxford University Press

A Re-entry as a Crucial Variable

Assume we have data on unemployment and health and unemployment for individuals in cells c in periods t. There are two types of individuals; those that loose health when they are in unemployment and those that do not. Denote the type of individual as $\Delta \in \{0, \delta\}$ and assume the share of individuals which experience a loss $\Delta = \delta$ is η. Assume also that this effect does not get worse across time and reverses completely if the individual finds a job again.[20] Denote the unemployment rate in group $\Delta \in \{0, \delta\}$ in cell c as $u_{ct}(\Delta)$. Assume that health of all individuals is 0 in employment. Assume that health of the employed is always 0. The average health in a cell c is then given by

$$\begin{aligned}
\bar{h}_{ct} &= u_{ct}(\delta)\,\eta \times \delta + u_{ct}(0)\,(1-\eta) \times 0 \\
&= \eta \times u_{ct}(\delta) \times \delta
\end{aligned}$$

which is simply the share of individuals that are affected by unemployment multiplied by their unemployment rate and the health damage of unemployment in this group. The true average effect of unemployment in this population is therefore:

$$\bar{h}^u - \bar{h}^e = \eta\delta.$$

Assume that unemployment dynamics in cell c at time t is driven by two exogenous probabilities, $\lambda_{ct}(\Delta)$ and $q_{ct}(\Delta)$, which are potentially a function of the type Δ. The parameter $\lambda_{ct}(\Delta)$ captures the likelihood that an individual of type Δ loses employment while $q_{ct}(\Delta)$ is the likelihood that an unemployed individual finds a job:

$$u_{ct}(\Delta) = \lambda_{ct}(\Delta)\,(1 - u_{ct-1}(\Delta)) + (1 - q_{ct}(\Delta)\,u_{ct-1}(\Delta)$$

Assuming stability of these parameters such a model would give long term unemployment of type Δ in cell c at time t as

$$u_{ct}(\Delta) = \frac{\lambda_{ct}(\Delta)}{\lambda_{ct}(\Delta) + q_{ct}(\Delta)}.$$

[20] The first assumption is not important. The second assumption is crucial and is discussed in detail in the text.

It is therefore impossible to capture the health effect of unemployment by comparing the unemployed to the employed population in equilibrium even if we control for cell characteristics. To see why note that comparing the employed to the unemployed in a cell c gives

$$\bar{h}^u_{ct} - \bar{h}^e_{ct} = \frac{u_{ct}(\delta)\,\eta}{u_{ct}(\delta)\,\eta + u_{ct}(0)\,(1-\eta)} \times \delta$$

where the first term is simply the share of individuals that suffer from unemployment among the unemployed in cell c at time t which in turn is a function of $q_{ct}(\delta)$ and $\lambda_{ct}(\delta)$. Assume for example, that those who lose health with unemployment were also most prone to lose their job, $\lambda_{ct}(\delta) > \lambda_{ct}(0)$. We will then get an overestimate of the effect of unemployment on health as, in the long run equilibrium, more individuals with health problems are in unemployment $u_{ct}(\delta) > u_{ct}(0)$.

There is now a large literature that uses job loss due to plant closures to get around this problem. This literature typically compares those who are unemployed due to an exogenous shock to the employed. Assume that we compare two groups; a group of employed individuals and a group who lost their job in the beginning of period t. The identifying assumption is that the share in those who lost their job due to plant closure is η and we therefore have

$$\bar{h}^{jobloss(t)}_{ct} - \bar{h}^e_{ct} = \eta\delta$$

which means the effect of job loss coincides with the effect of unemployment in period t. However, in the next period a share of the affected individuals $q_{ct}(\delta)$ has found a job. Comparing the health of individuals that have lost their job in the previous period with that that are in employment therefore yields

$$\bar{h}^{jobloss(t)}_{ct+1} - \bar{h}^e_{ct+1} = (1 - q_{ct}(\delta))\,\eta\delta < \eta\delta$$

in other words, the health effect from a job loss is now smaller than the health effect of unemployment. How large this deviation is depends on the probability of re-entry $q_{ct}(\delta)$. Only if re-entry is shut off, $q_c(\delta) = 0$, the effect of unemployment can be identified by comparing individuals who lost their job to individuals in employment.[21]

[21] Note that it is not a solution to condition on unemployment after job-loss which would yield

$$\bar{h}^{u,jobloss(t)}_{ct+1} - \bar{h}^e_{ct+1} = \frac{1 - q_{ct}(\delta)\,\eta}{(1 - q_{ct}(\delta))\,\eta + (1 - q_{ct}(0))\,(1-\eta)}\delta$$

as conditioning on unemployment still allows individuals to exit unemployment at different rates.

We follow the basic idea of plant closure, exogenous job loss, but instrument for unemployment at the cell level. This has the disadvantage that we cannot condition directly on unemployment and employment of individuals within cells. Instead, we use a difference-in-difference strategy. In the framework here this is equivalent to a comparison of health in the same cell across time as a function of unemployment. Formally this measure is given by

$$\frac{\bar{h}_{ct+1} - \bar{h}_{ct}}{\hat{u}_{ct+1} - \hat{u}_{ct}} = \frac{\hat{u}_{ct+1}(\delta)\,\eta \times \delta - \hat{u}_{ct}(\delta)\,\eta \times \delta}{\hat{u}_{ct+1}(\delta)\,\eta + \hat{u}_{ct+1}(0)(1-\eta) - \hat{u}_{ct}(\delta)\,\eta - \hat{u}_{ct}(0)(1-\eta)}.$$

where $\hat{u}_{ct+1}$ and $\hat{u}_{ct}$ are the fitted values from our first stage. The main change in unemployment captured by these fitted values is the dramatic rise in unemployment after 2007 which was both driven by an abrupt decrease in λ_{ct} and q_{ct} for individuals of both types $\Delta \in \{0, \delta\}$. The identifying assumption we make is that the exogenous changes in employment opportunities did not affect unemployment differently in the two types, i.e. we assume that

$$\begin{aligned} \hat{u}_{ct+1}(\delta) - \hat{u}_{ct}(\delta) &= \hat{u}_{ct+1}(0) - \hat{u}_{ct}(0) \\ &= \hat{\beta} \times (construcIV_{c,t+1} - construcIV_{c,t}) \end{aligned} \tag{12}$$

so that we can then identify the true effect of unemployment on health by comparing cells with large changes in unemployment and cells with small changes in unemployment. Formally our IV estimate in the second stage then captures the true effect of unemployment of crime

$$\frac{\bar{h}_{ct+1} - \bar{h}_{ct}}{\hat{u}_{ct+1} - \hat{u}_{ct}} = \eta\delta.$$

As should be clear from the previous discussion our identifying assumption in equation (12) in fact consists of two assumptions. First, the change in job loss was the same for individuals of both types Δ. For this we need to find a way to exploit the different exposure of cells to the exogenous Macro shock that hit the Spanish economy in the years following 2007. Our aim is to have a shock similar to plant closures where job loss did not discriminate between workers. Second, the likelihood of escaping unemployment needs to be the same for both types. It is here where the particularities of Spanish case provides a unique setting as re-entry into employment was almost impossible in some sectors.

Table 1: Summary Statistics

general questions (2001, 2003, 2006, 2011)

Variable	Obs	Mean	Std. Dev
age	46247	39.262	11.137
male	46247	0.583	0.493
secondary education	46247	0.559	0.496
college education	46247	0.220	0.414
unemployed	46247	0.162	0.369
reported good health	46247	0.789	0.408
mental disorder diagnosed by doctor	46247	0.068	0.251

health questions (2006, 2011)

Variable	Obs	Mean	Std. Dev
reported good health	24856	0.790	0.407
mental disorder diagnosed by doctor	24856	0.084	0.278
mental disorder reported	24856	0.087	0.283
mental disorder GHQ questionaire average score	24856	0.108	0.192
chronic backpain diagnosed	24856	0.224	0.417
chronic headache diagnosed	24856	0.093	0.290
heart attack diagnosed	24856	0.029	0.167
stroke diagnosed	24856	0.004	0.063
present or previous employment in construction	24856	0.109	0.311

GHQ Mental Health Survery (2006, 2011)

Variable	Obs	Mean	Std. Dev
lost sleep	24856	0.240	0.427
felt under strain	24856	0.257	0.437
unable to enjoy activities	24856	0.102	0.303
unhappy or depressed	24856	0.141	0.348
feeling unhappy	24856	0.059	0.236
unable to concentrate	24856	0.114	0.318
unable to make decisions	24856	0.042	0.201
unable to overcome difficulties	24856	0.120	0.325
unable to face problems	24856	0.056	0.229
feeling useless	24856	0.064	0.245
lost self confidence	24856	0.065	0.246
worthless person	24856	0.032	0.177

Table 2: Unemployment and Health (OLS)

VARIABLES	(1) reported good health	(2) reported good health	(3) mental disorder diagnosed by doctor	(4) mental disorder diagnosed by doctor	(5) chronic backpain diagnosed by doctor	(6) chronic headache diagnosed by doctor	(7) heart disease diagnosed by doctor	(8) stroke diagnosed by doctor
unemployed	-0.205***	-0.203***	0.165***	0.163***	0.0265	0.0426*	0.0557**	0.0454**
	(0.0204)	(0.0204)	(0.0198)	(0.0232)	(0.0213)	(0.0218)	(0.0227)	(0.0229)
male	0.0613***							
	(0.00557)							
under 40	0.123***							
	(0.00554)							
province and year fixed effects	yes	no	no	no	no	no	no	no
cell and year fixed effects	no	yes	yes	yes	yes	yes	yes	yes
survey years	2001-11	2001-11	2001-11	2006-11	2006-11	2006-11	2006-11	2006-11
Observations	46,330	46,330	46,358	25,544	25,544	25,544	25,544	25,544
R-squared	0.041	0.047	0.053	0.049	0.047	0.061	0.050	0.019

Robust standard errors clustered at the province level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All dependent variables are weighted by their standard deviation. Cells are formed by all possible interactions between a dummy for male, a dummy for under 40 and 51 province dummies (2x2x51 = 204 cells).

Table 3: Construction Sector Employment as Predictor of Unemployment

VARIABLES	(1) unemployed	(2) unemployed	(3) unemployed	(4) unemployed	(5) unemployed	(6) unemployed	(7) unemployed
employment growth	-2.571***						
	(0.603)						
employment growth (construction)		-2.340***	-2.258***	-2.325***	-2.429***	-2.454***	-16.48
		(0.582)	(0.451)	(0.395)	(0.420)	(0.456)	(11.08)
employment growth (not construction)		-0.278					
		(0.968)					
cell fixed effects	yes	yes	yes	yes	yes	yes	yes
year fixed effects	yes	yes	yes	yes	yes	yes	yes
survey years	2001-11	2001-11	2001-11	2001-11	2001-11	2006-11	2001-2006
Observations	46,358	46,358	46,358	46,275	46,358	25,544	35,579
R-squared	0.067	0.067	0.067	0.085	0.081	0.068	0.044

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Sample are the years 2001 to 2011 except for column (6) where the sample is only 2006 and 2011 and column (7) where the sample is 2001 to 2006. Regressions in columns (1), (2), (3) , (6) and (7) use cells defined by provinces, sex and a dummy of age<40. Column (4) uses two age dummies <30, >50. Column (5) instead adds a dummy for college education.

Table 4: Main Results - Mental Health and Unemployment (IV)

VARIABLES	(1) reported good heatlh	(2) mental disorder diagnosed by doctor	(3) mental disorder reported	(4) summary score GHQ12 mental health survery	(5) In the last couple of weeks have you… Lost much sleep over worry?	(6) Felt constantly under strain?	(7) Been able to enjoy your normal day-to-day activities?	(8) Been feeling unhappy and depressed?
unempl	-0.741**	1.103**	1.169**	0.911*	0.447	1.337***	0.370	0.686
	(0.364)	(0.498)	(0.498)	(0.470)	(0.348)	(0.496)	(0.368)	(0.454)
Observations	25,544	25,544	25,544	24,914	25,082	25,055	25,061	25,060
R-squared	0.010			0.030	0.041		0.028	0.027

VARIABLES	(9) In the last couple of weeks have you… Been feeling reasonably happy, all things considered?	(10) Been able to concentrate on whatever you are doing?	(11) Felt capable of making decisions about things?	(12) Felt that you couldn't overcome your difficulties?	(13) Been able to face up to your problems?	(14) Felt that you were playing a useful part in things?	(15) Been losing self-confidence in yourself?	(16) Been thinking of yourself as a worthless person?
unempl	0.572	0.859*	0.334	0.772*	0.238	0.886**	0.255	0.150
	(0.477)	(0.446)	(0.409)	(0.399)	(0.482)	(0.370)	(0.504)	(0.535)
Observations	25,059	25,102	25,075	25,063	25,046	25,035	25,045	25,025
R-squared	0.027		0.026	0.020	0.024	0.020	0.034	0.031

Robust standard errors clustered at the cell level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All dependent variables are weighted by their standard deviation. In columns (4) to (16) higher values are always more negative outcomes. Variables are recoded such that they take values 0 (better and as usual) and 1 (worse than usual). The summary scores is the average score divided by 12. All regressions control for cell and year fixed effects. Cells are defined by provinces, sex and a dummy of age<40.

Table 5: Other Health Outcomes

VARIABLES	(1) chronic backpain diagnosed by doctor	(2) chronic headache diagnosed by doctor	(3) heart disease diagnosed by doctor	(4) stroke diagnosed by doctor	(5) takes medicines
unemployed	0.659	0.873*	0.519	-0.157	1.005**
	(0.551)	(0.505)	(0.319)	(0.301)	(0.469)
cell fixed effects	yes	yes	yes	yes	yes
year fixed effects	yes	yes	yes	yes	yes
Observations	25,544	25,544	25,544	25,544	25,544
R-squared	0.005			0.004	

Robust standard errors clustered at the cell level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All dependent variables are weighted by their standard deviation. Cells are defined by provinces, sex and a dummy of age<40.

Table 6: Unemployment and Suicides (IV)

	(1)	(2)	(3)	(4)	(5)
VARIABLES			ln(suicides)		
unemployment rate	-0.303	4.231**	4.706**	4.979**	3.846**
	(0.808)	(1.981)	(2.008)	(2.122)	(1.624)
cell fixed effects	yes	yes	yes	yes	yes
year fixed effects	yes	yes	yes	yes	yes
control of ln(population)	no	no	yes	yes	yes
province time trend	no	no	no	yes	no
survey years	2001-11	2007-11	2007-11	2007-11	2007-11
Observations	2,035	921	921	921	1,283
R-squared	0.944	0.953	0.954	0.959	0.943

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Unmployment rate is the unemployment rate at the cell level. Column (1) uses data from 2001 till 2011. All other columns use data from 2007 to 2011. Columns (1) to (4) use cells defined by provinces, sex and a dummy of age<40. Column (5) uses two age dummies <30, >50.

Table 7: Mental Health and Unemployment (IV), 3 age categories

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
						In the last couple of weeks have you…		
VARIABLES	reported good heatlh	mental disorder diagnosed by doctor	mental disorder reported	summary score GHQ12 mental health survery	Lost much sleep over worry?	Felt constantly under strain?	Been able to enjoy your normal day-to-day activities?	Been feeling unhappy and depressed?
unempl	-0.646*	1.044**	1.083***	0.828*	0.371	1.192**	0.324	0.811*
	(0.337)	(0.426)	(0.398)	(0.446)	(0.350)	(0.482)	(0.375)	(0.436)
Observations	25,544	25,544	25,544	24,914	25,082	25,055	25,061	25,060
R-squared	0.010			0.030	0.041		0.028	0.027

	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
				In the last couple of weeks have you…				
VARIABLES	Been feeling reasonably happy, all things considered?	Been able to concentrate on whatever you are doing?	Felt capable of making decisions about things?	Felt that you couldn't overcome your difficulties?	Been able to face up to your problems?	Felt that you were playing a useful part in things?	Been losing self-confidence in yourself?	Been thinking of yourself as a worthless person?
unempl	0.519	0.782*	0.279	0.660	0.232	0.861**	0.143	0.104
	(0.412)	(0.418)	(0.379)	(0.407)	(0.438)	(0.367)	(0.434)	(0.478)
Observations	25,059	25,102	25,075	25,063	25,046	25,035	25,045	25,025
R-squared	0.027		0.026	0.020	0.024	0.020	0.034	0.031

Robust standard errors clustered at the cell level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All dependent variables are weighted by their standard deviation. In columns (4) to (16) higher values are always more negative outcomes. Variables are recoded such that they take values 0 (better and as usual) and 1 (worse than usual). The summary scores is the average score divided by 12. All regressions control for cell and year fixed effects. Cells are defined by provinces, sex a dummy for age<30 and a dummy for age>50.

Table 8: Mental Health and Unemployment (IV), dummy for college

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
						In the last couple of weeks have you…		
VARIABLES	reported good heatlh	mental disorder diagnosed by doctor	mental disorder reported	summary score GHQ12 mental health survery	Lost much sleep over worry?	Felt constantly under strain?	Been able to enjoy your normal day-to-day activities?	Been feeling unhappy and depressed?
unempl	-0.432	0.923**	0.974**	0.855**	0.396	1.248***	0.401	0.658*
	(0.345)	(0.391)	(0.394)	(0.389)	(0.325)	(0.411)	(0.297)	(0.381)
Observations	25,461	25,461	25,461	24,856	25,021	24,994	25,000	25,000
R-squared	0.061			0.047	0.053		0.036	0.041

	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
				In the last couple of weeks have you…				
VARIABLES	Been feeling reasonably happy, all things considered?	Been able to concentrate on whatever you are doing?	Felt capable of making decisions about things?	Felt that you couldn't overcome your difficulties?	Been able to face up to your problems?	Felt that you were playing a useful part in things?	Been losing self-confidence in yourself?	Been thinking of yourself as a worthless person?
unempl	0.480	0.710*	0.365	0.680**	0.333	0.777**	0.143	0.178
	(0.391)	(0.381)	(0.387)	(0.343)	(0.387)	(0.317)	(0.394)	(0.446)
Observations	24,999	25,041	25,014	25,002	24,987	24,974	24,985	24,966
R-squared	0.041	0.000	0.035	0.040	0.033	0.042	0.041	0.040

Robust standard errors clustered at the cell level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All dependent variables are weighted by their standard deviation. In columns (4) to (16) higher values are always more negative outcomes. Variables are recoded such that they take values 0 (better and as usual) and 1 (worse than usual). The summary scores is the average score divided by 12. All regressions control for cell and year fixed effects. Cells are defined by provinces, sex a dummy for age<40 and a dummy for college education.

Table 9: Robustness

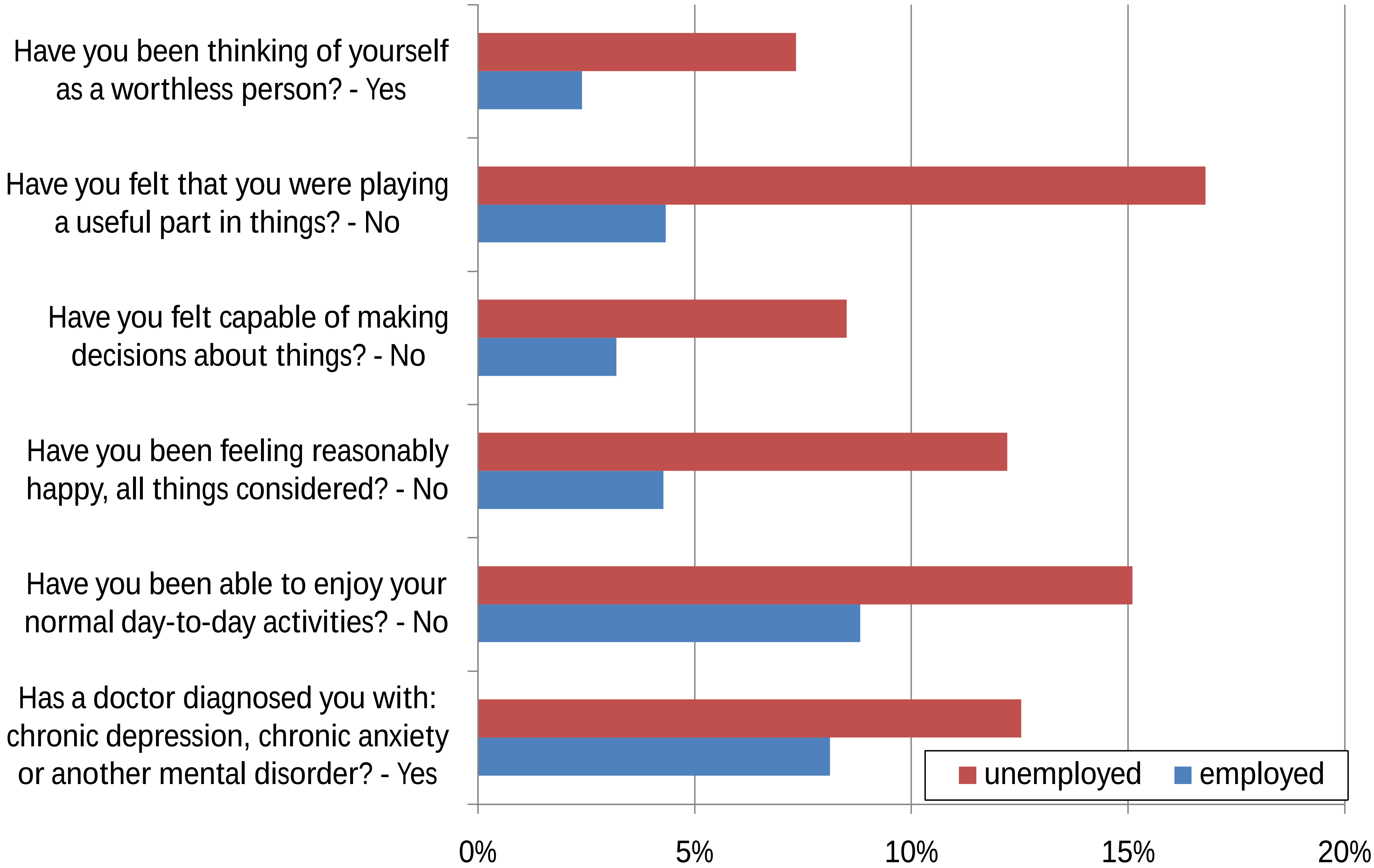
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	total employment as IV	growth in last three years as IV	construction employment at province level as IV	average unemployment at cell level	including inactive population	including early waves	province time trends	male time trend
VARIABLES	mental disorder diagnosed by doctor							
unemployed	1.634**	1.103**	0.415	1.880**	1.139**	0.745*	0.853*	1.749***
	(0.648)	(0.498)	(1.468)	(0.780)	(0.571)	(0.442)	(0.441)	(0.585)
cell and year fixed effects	yes	yes	no	yes	yes	yes	yes	yes
male dummy, under 40 dummy, province and year fixed effects	no	no	yes	no	no	no	no	no
province-specific time trend	no	no	no	no	no	no	yes	no
male-specific time trend	no	no	no	no	no	no	no	yes
Observations	25,544	25,544	25,544	25,544	36,563	46,358	46,358	46,358
R-squared			0.028	0.044		0.001		

Robust standard errors clustered at the cell level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All dependent variables are weighted by their standard deviation. Cells are defined by provinces, sex and a dummy for age<40. Columns (1) to (4) include years 2006 and 2011. Columns (5) to (7) include years 2001, 2003, 2006 and 2011. Column (3) uses unemployment at the cell level from the EPA instead of the dummy for unemployment from the NHS.

Table A1: Diagnosed Mental Disorders in 2006 and 2011

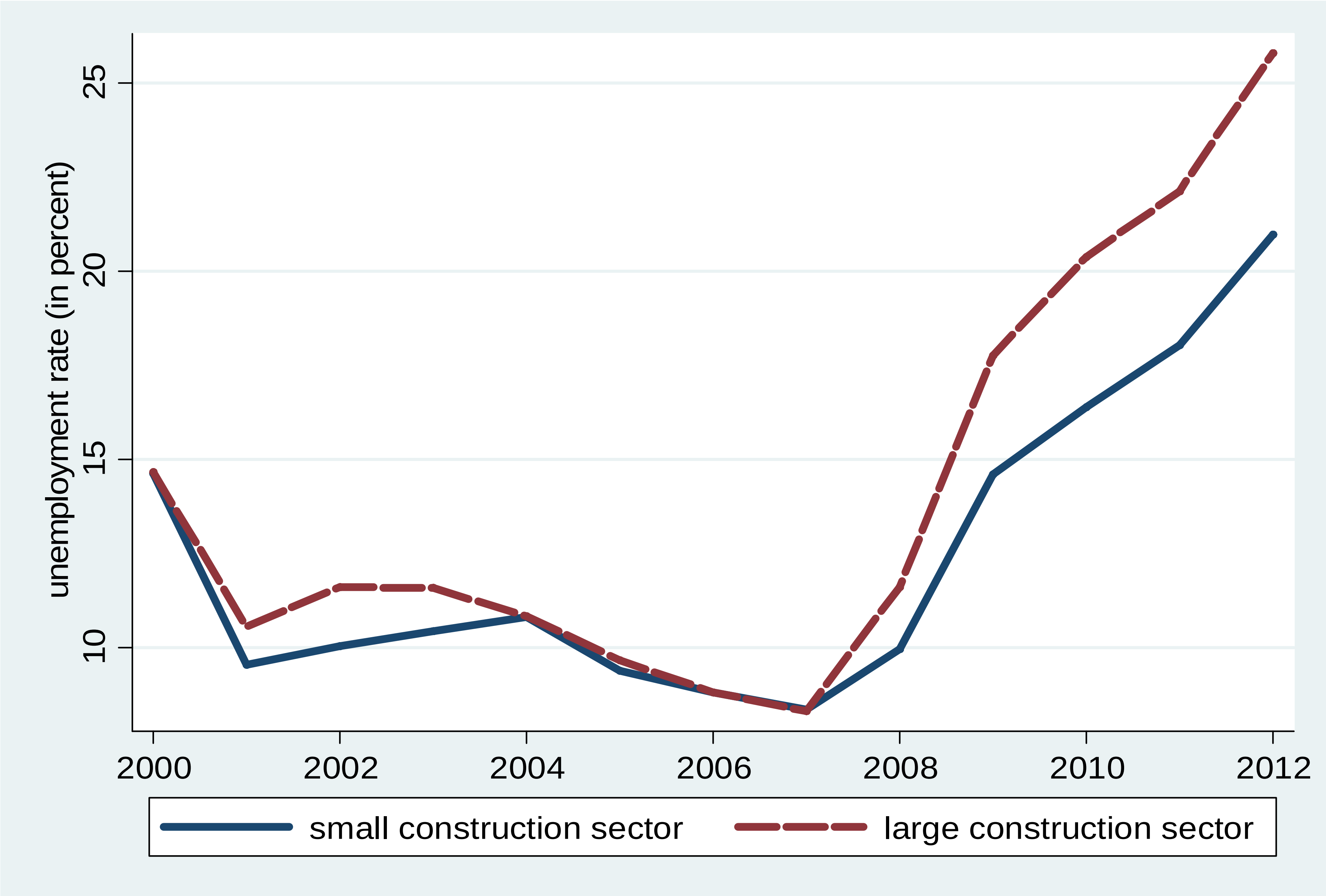
mental disorder diagnosed by doctor	2006	2011	difference
all individuals	0.098	0.072	-0.025
present or previous employment in construction	0.052	0.053	0.001
present or previous employment not construction	0.104	0.075	-0.029

Figure 1: Unemployment and Mental Health



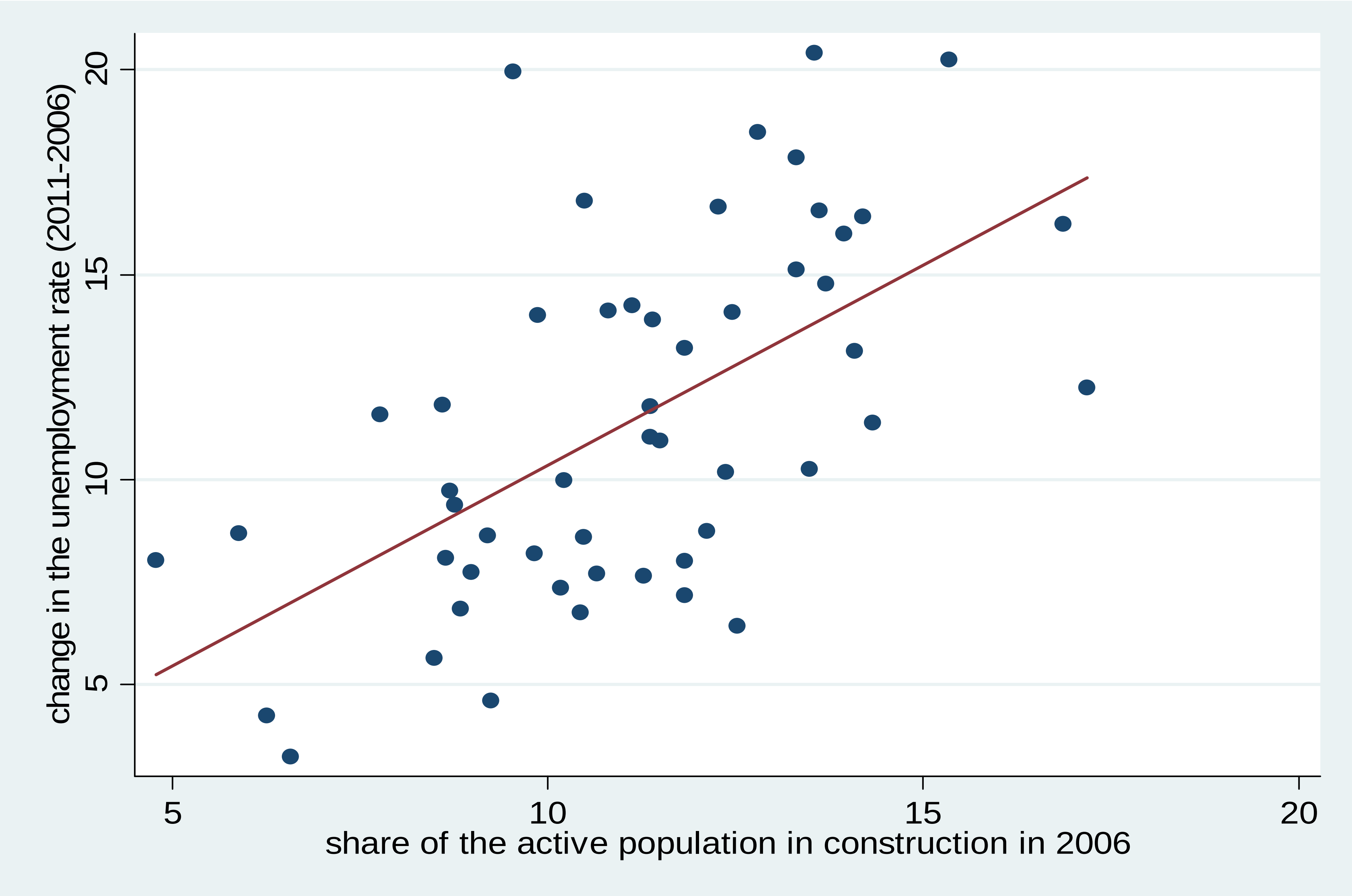
Note: Yes and No respectively represent "worse than usual" answers.
Source: Spanish National Health Survey. Years 2006 and 2011.

Figure 2: Spanish Unemployment in the Financial Crisis



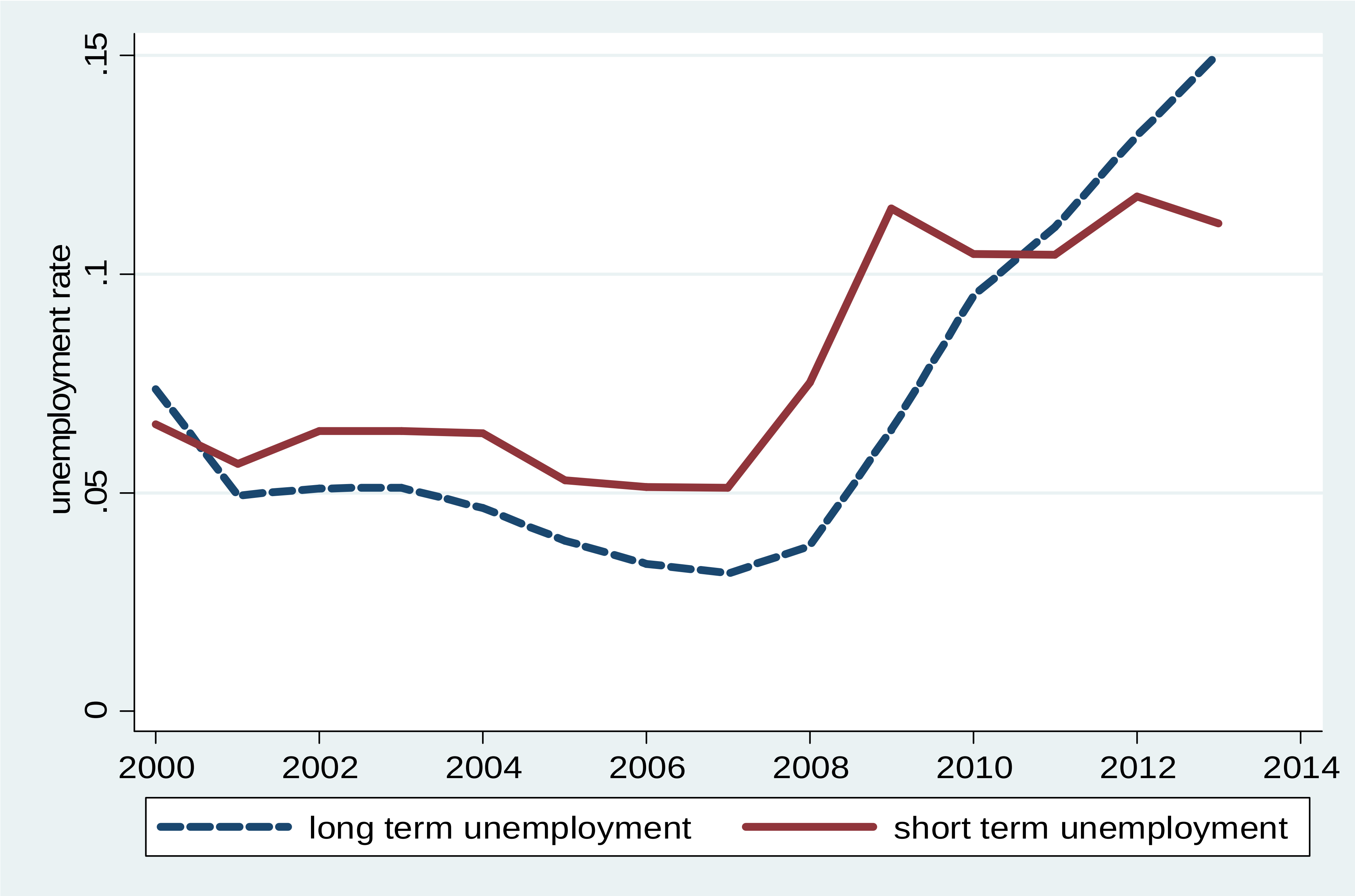
Source: Spanish Labor Force Survey

Figure 3: Changes in Unemployment and the Construction Sector



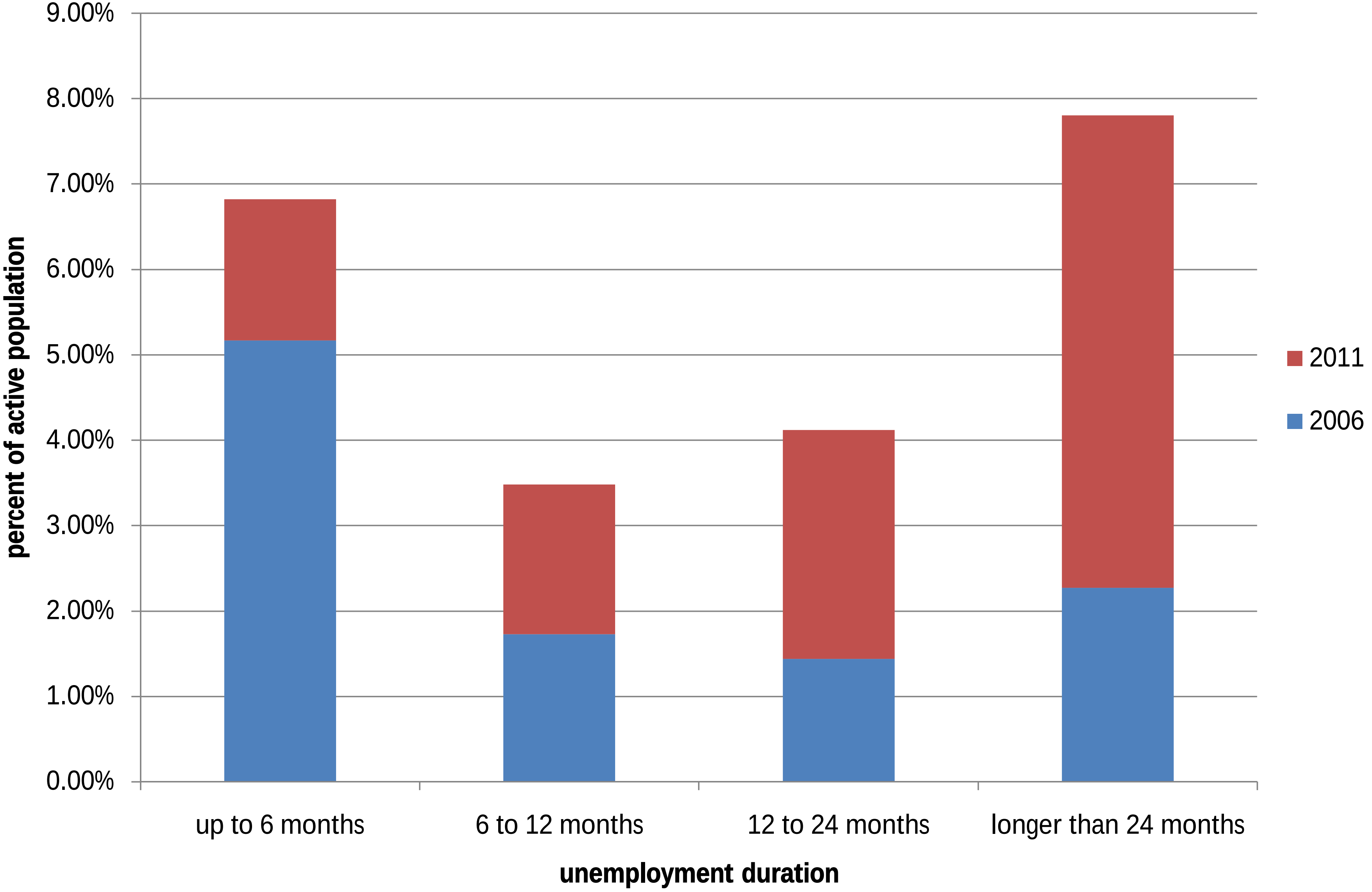
Source: Spanish Labor Force Survey

Figure 4: Short- and Long-Term Unemployment



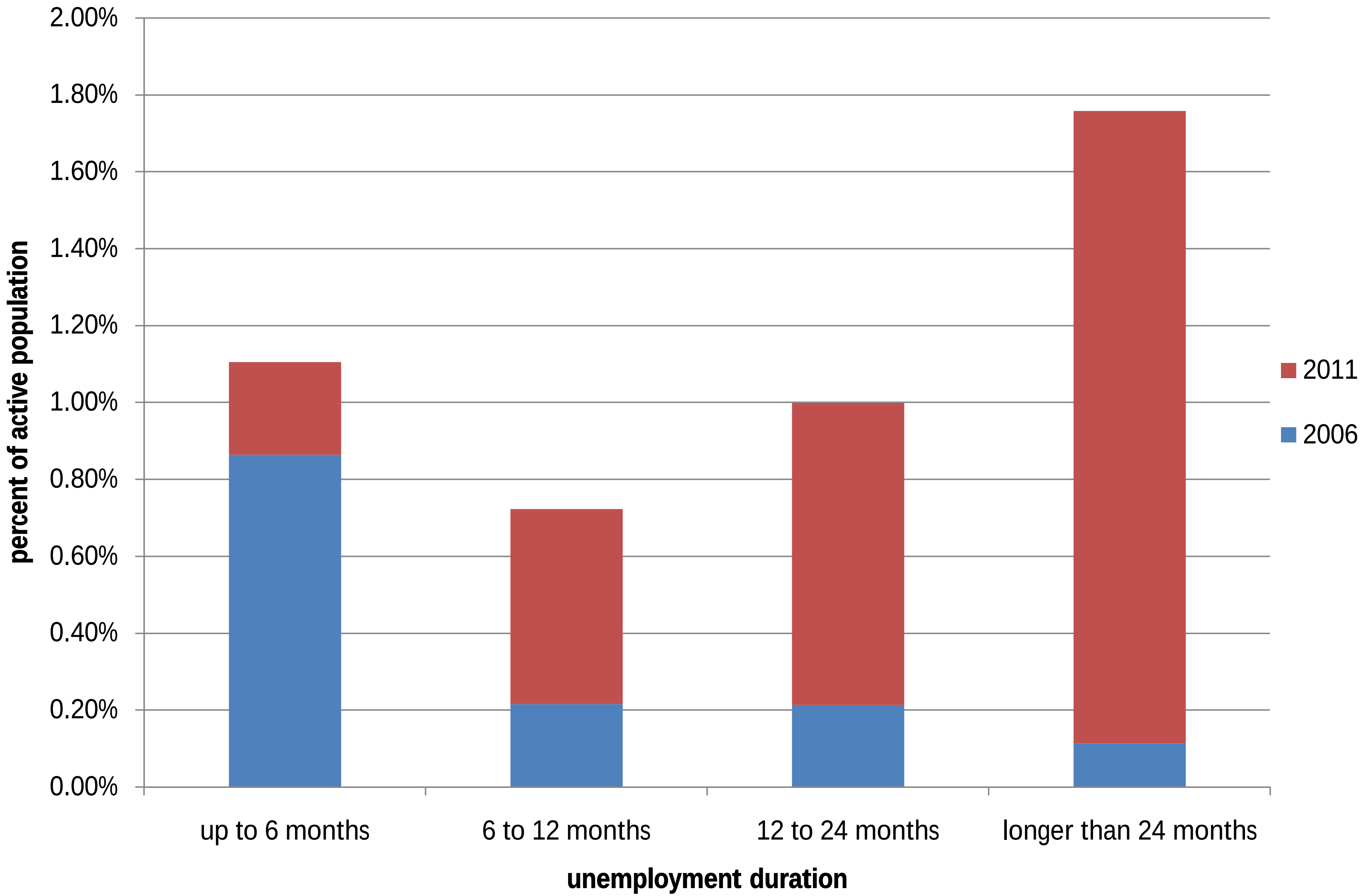
Source: Spanish Labor Force Survey

Figure 5a: Increase in Unemployment Duration (All Individuals)



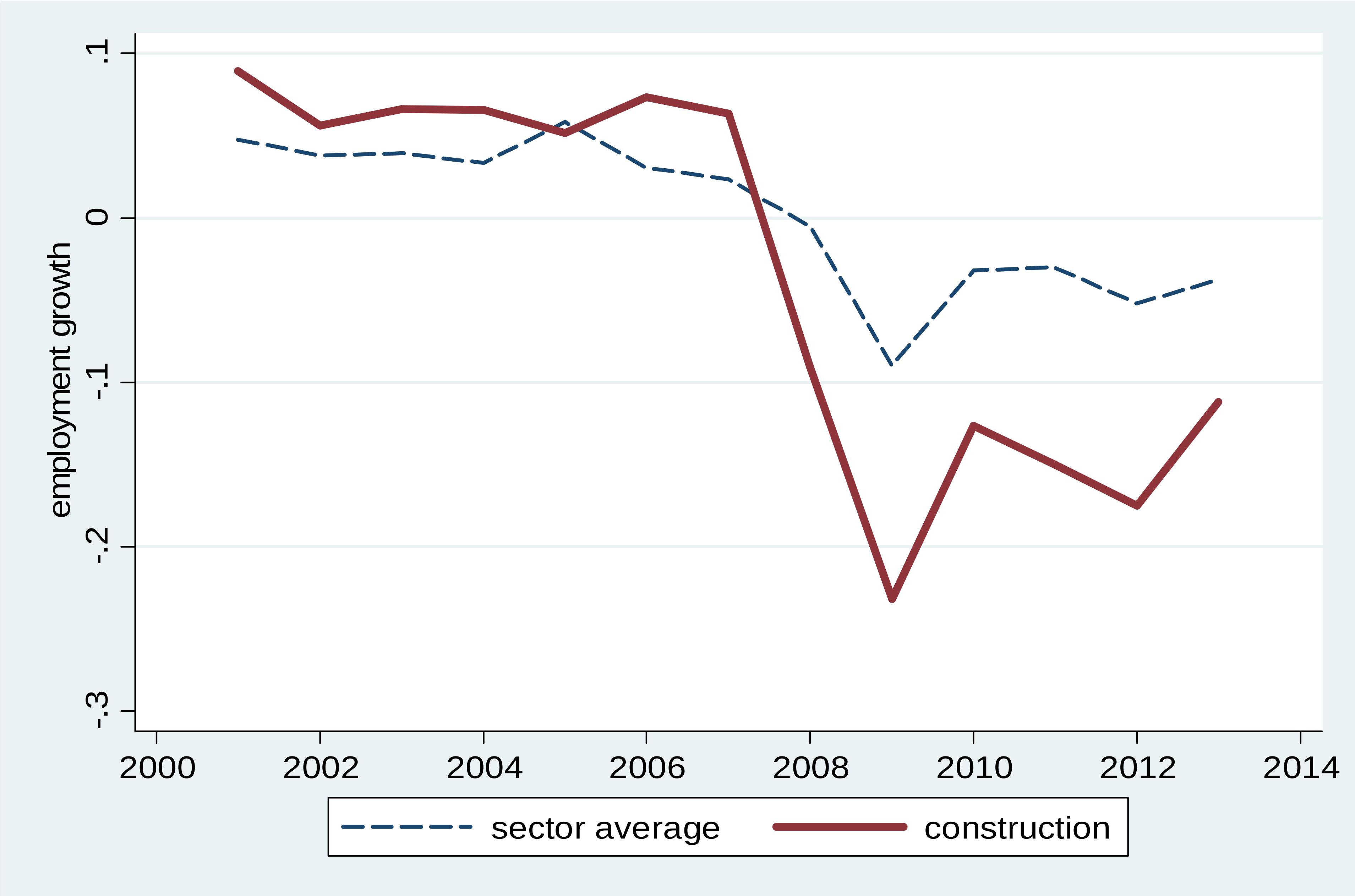
Source: Spanish National Health Survey

Figure 5b: Increase Unemployment Duration (Formerly Employed in Construction)



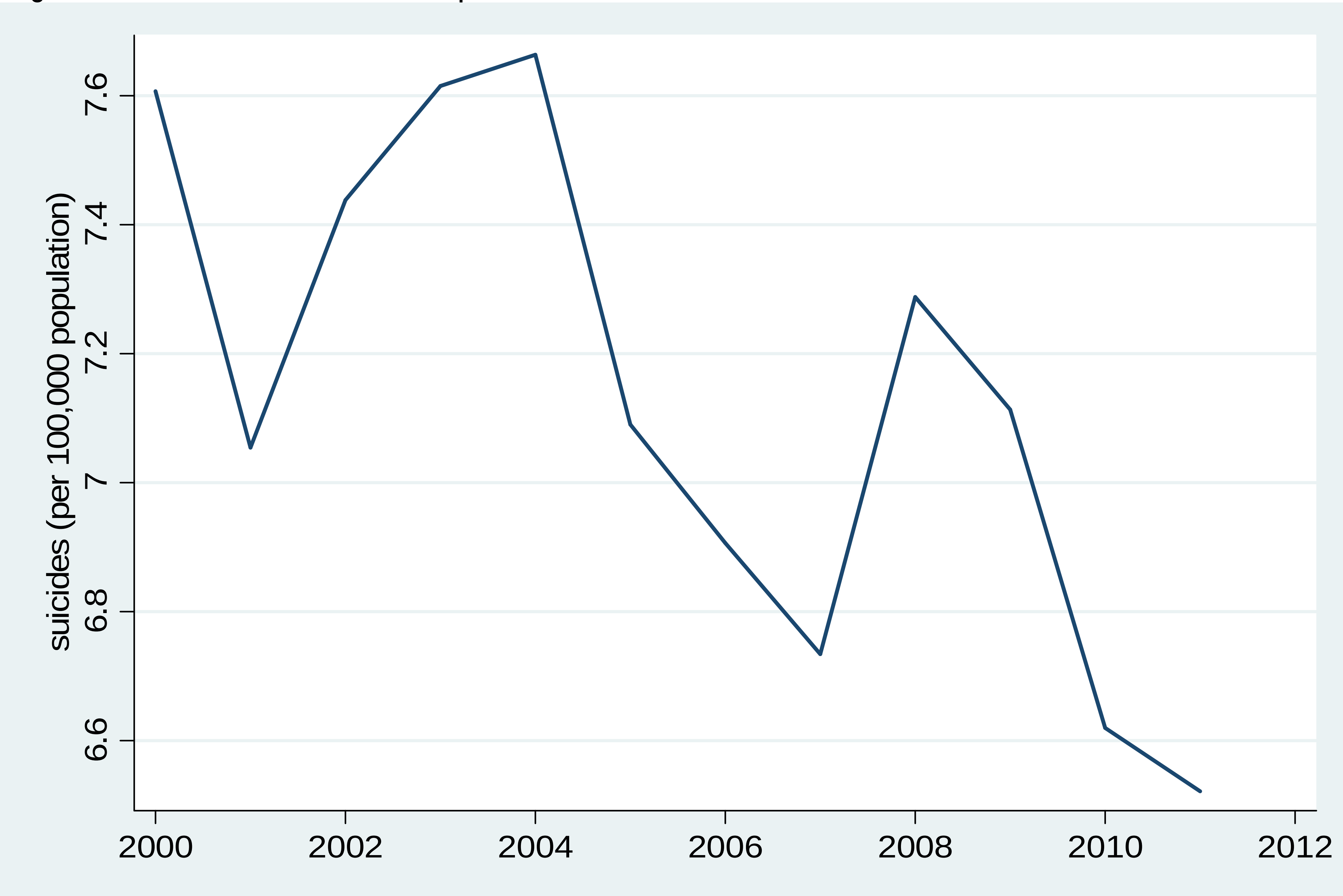
Source: Spanish National Health Survey

Figure 6: Employment Growth in Spain



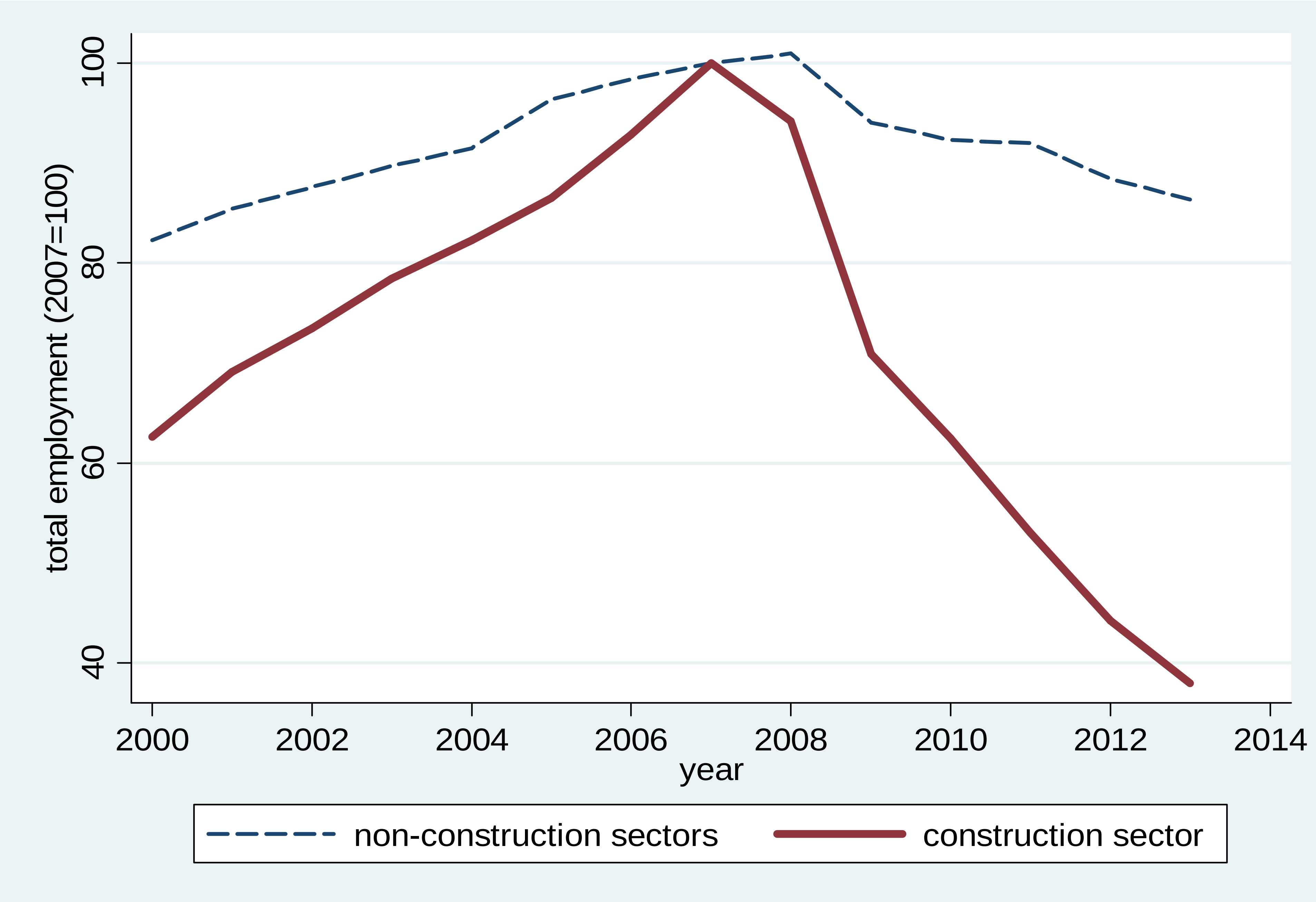
Source: Spanish Labor Force Survey

Figure 7: Number of Suicides in Spain



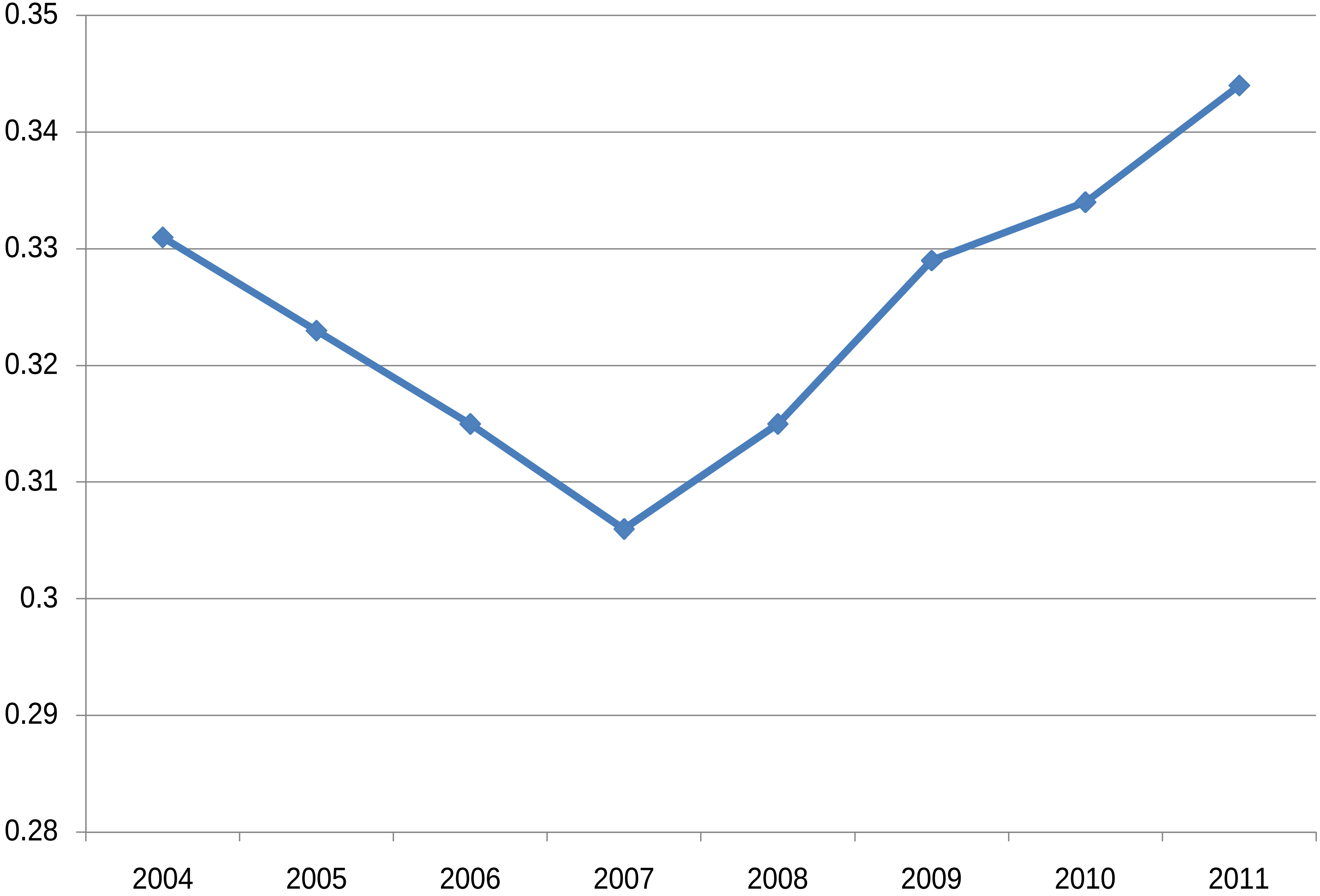
Source: Spanish Statistical Office. Death reports.

Figure A1: Boom and Bust of Employment in the Construction Sector



Source: Spanish Labor Force Survery

Figure A2: The Gini Coefficient in Spain



Source: OECD Income Distribution and Poverty Database

Figure A3: Fitted Unemployment Rates in 2006 and 2011

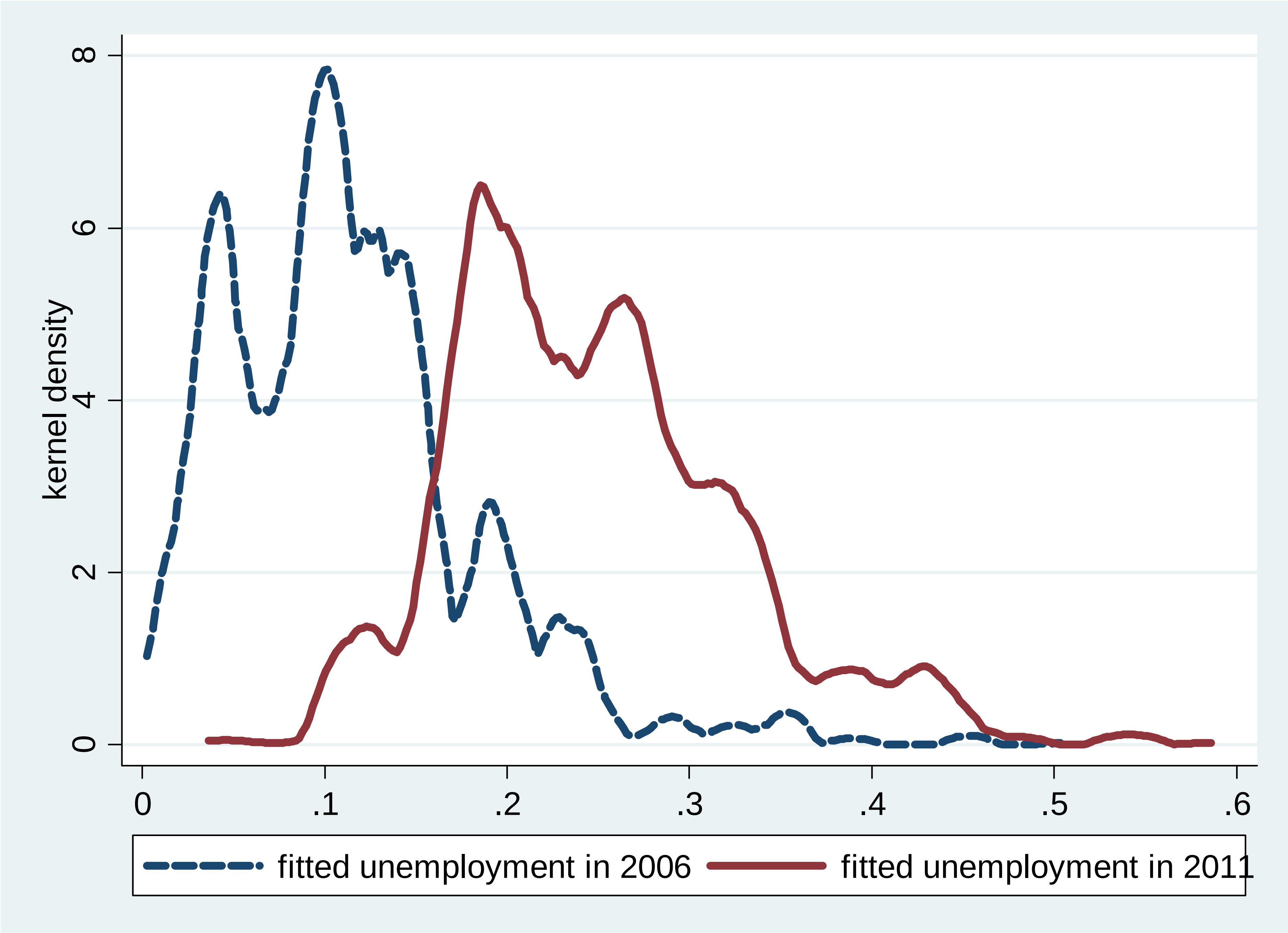
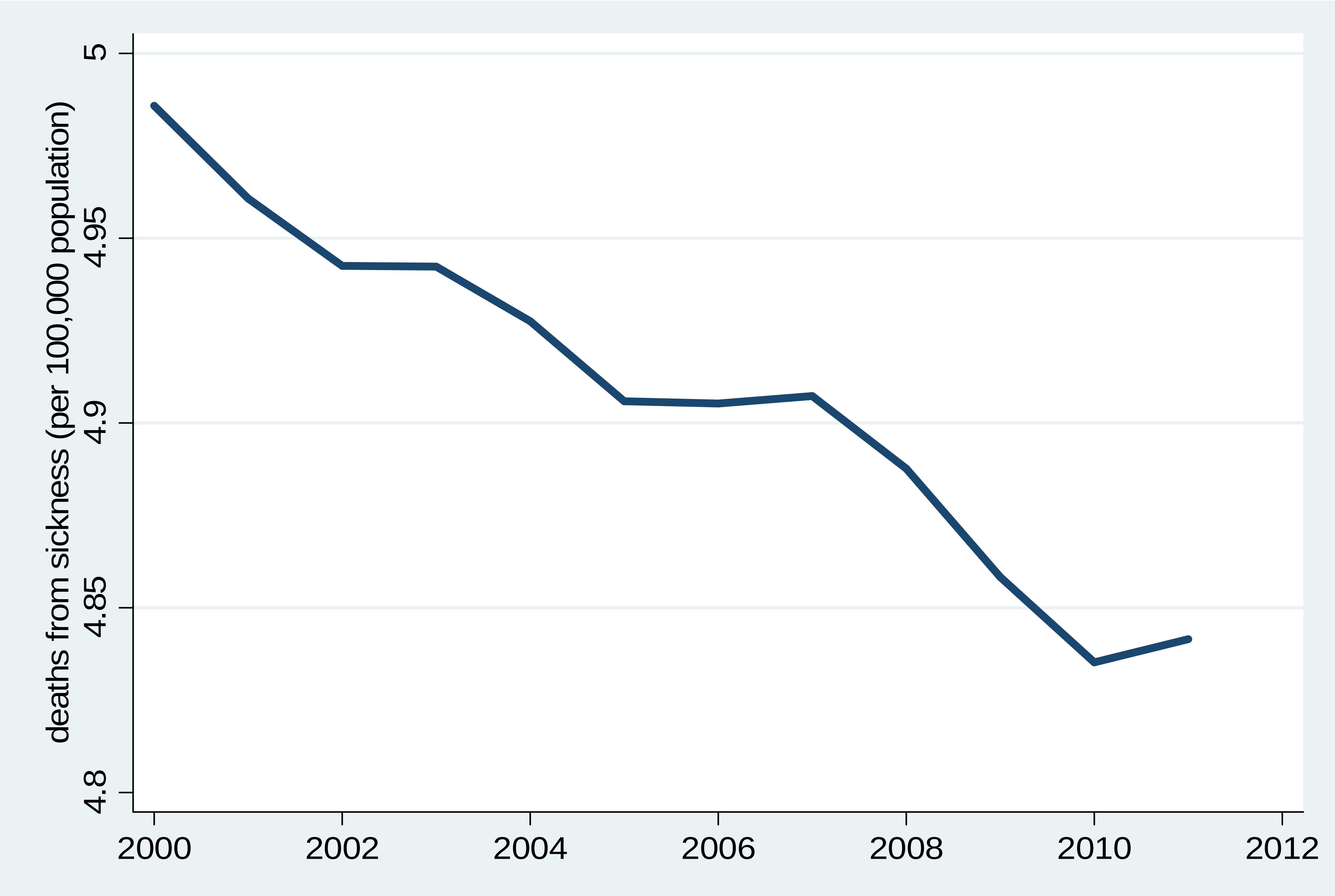


Figure A4: Death Rate from Sickness in Spain



Note: Figure shows the sum of deaths from the four main illnesses (cancer, respiratory, infectious and cardiovascular deseases).