Working papers series

WP ECON 18.11

The social cost of unemployment: the Spanish labour market from a social welfare approach

Lucía Gorjón
ISEAK & FEDEA

Sara de la Rica
Universidad del País Vasco & FEDEA

Antonio Villar
Universidad Pablo de Olavide & Ivie

Keywords: social cost of unemployment, unemployment benefits, incidence, severity and hysteresis of unemployment, Spanish labour market, regional differences.

JEL Classification: J64, J65, I31.

Department of Economics

The social cost of unemployment: the Spanish labour market from a social welfare approach

Lucía Gorjón

ISEAK & FEDEA

Sara de la Rica

Universidad del País Vasco & FEDEA

Antonio Villar

Universidad Pablo de Olavide & Ivie

Abstract

This paper proposes a protocol for considering the social cost of unemployment by taking into account three different aspects: incidence, severity and hysteresis. Incidence refers to the conventional unemployment rate; severity takes in both unemployment duration and the associated income loss; and hysteresis refers to the probability of remaining unemployed. The social cost of unemployment is regarded as a welfare loss, which is measured by a utilitarian social welfare function whose arguments are the individual disutilities of unemployed workers. Each individual disutility is modelled as a function of income loss, unemployment duration and hysteresis. The resulting formula is simple and easy to understand and implement. We apply this assessment protocol to the Spanish labour market, using the official register of unemployed workers compiled by the Public Employment Service.

JEL Classification: J64, J65, I31

Key words: social cost of unemployment, unemployment benefits, incidence, severity and hysteresis of unemployment, Spanish labour market, regional differences.

Acknowledgements: This work has benefited from comments and suggestions on an earlier draft made by Samuel Bentolila, Juan José Dolado and José I. García Pérez. The first and second authors acknowledge financial support from the Spanish Ministry of Education and Science, ECO2015-67105-R. The third author acknowledges financial support from Project ECO2015-65408-R (MINECO/FEDER/UE).

1 Introduction

"What you measure affects what you do.
If you don't measure the right thing, you don't do the right thing."

Joseph Stiglitz

The labour market has suffered a massive shock with the global financial crisis and unemployment rates in many countries have rocketed to levels not seen for decades. Most economies are now recovering their pre-crisis levels of activity and unemployment rates are consistently declining. Note, however, that aggregate unemployment rates may hide the presence of large asymmetries in the labour market, because unemployment has hit different geographical areas and types of workers differently, even within the same country. Those asymmetries involve not only differences in its *incidence* (unemployment rates) but also in *severity* (unemployment duration and benefits received), and *hysteresis* (the likelihood of remaining unemployed).

To illustrate this point we look at the Spanish labour market. Spain is one of the countries hardest hit by the crisis, with more than one fifth of the active population being unemployed for a long while. The cycle has recently changed and Spain now exhibits high rates of growth and substantial reductions in unemployment, especially among those with shorter unemployment durations. Long-term unemployment, though, is much less sensitive to the recovery and the data show that reductions in unemployment go together with an increase in the average length of unemployment for the long-term unemployed. Moreover, the proportion of unemployed workers with no access to unemployment benefits has also increased. This points to the formation of a progressively marginalised group of workers who will find it extremely hard to find a job (see Bentolila & García-Pérez (2017) for a discussion in terms of survival rates). It is also clear that unemployment rates are far from capturing the social cost of unemployment.

Our starting point is the idea that to measure the social cost of unemployment one must first take properly into account all three aspects involved (incidence, severity and hysteresis) and secondly consider the differences between different Spanish regions. We propose to deal with this assessment problem in terms of a social welfare function that captures the welfare loss to society derived from the disutility of the unemployed. The rationale of adopting a social welfare perspective is quite straightforward: unemployment entails a welfare loss for society and thus it is sensible to compute the size of that loss and not only the incidence of unemployment. This is the type of approach pioneered by Dalton (1920) in the analysis of inequality, later enhanced and perfected by Atkinson (1970) and Sen (1973) among many others. It also shares the spirit of those poverty indices that combine incidence and intensity measures (see Chakarvarty (2009), Villar (2017) for a discussion and detailed references). Indeed, there are contributions that have used the standard approach to poverty measurement to incorporate duration in the assessment of unemployment (see Sengupta (2009), Shorrocks (2009 a, b)); Goerlich & Miñano (2018) provide an application of this methodology to the Spanish labour market.

Rather than starting from an axiomatically based aggregate indicator like those mentioned above, our assessment function is obtained from the aggregation of individual disutility levels.[1] We model the individual agent's disutility on being unemployed at a given point in time as a function of income loss, unemployment duration and the probability of remaining unemployed. Our approach takes explicitly into account the different levels of severity of unemployment depending on whether there is access to unemployment benefits or social subsidies and the duration of unemployment. We also allow for a non-linear impact of unemployment duration on disutility, as the long-term unemployed suffer not only from an accumulation of low income periods but also from a reduction in the probability of exiting their status and from a whole array of personal and social difficulties that affect self-respect, social involvement and social inclusion.[2] We associate the degree of convexity of duration with the probability of remaining unemployed.

The social cost of unemployment is obtained by aggregating the disutility of unemployed individuals; it results in a function that involves the number of unemployed people, unemployment spells, transition probabilities and income loss (the difference between the market wage and unemployment benefit, if any, for each unemployed worker). The resulting assessment function is simple, easy to interpret, based on an explicit model, and applicable to real-life problems.

[1] There is some parallelism with the paper by Jones & Klenow (2016) in the use of a micro approach to address the problem and a multidimensional indicator that goes beyond the unemployment rate (beyond the GDP in their case).

[2] See Winter-Ebmer (2016) and de la Rica and Gorjón (2017) for a discussion. Recall that the United Nations have for many years been using the rate of long-term unemployment as a proxy for (lack of) social inclusion.

The rest of the paper is organised as follows. Section 2 presents the basic model and the assessment formula, which consists of the overall social disutility. This is obtained by aggregating the disutility of individuals, which is derived from a simple utility maximisation program. Individual disutility is a convex function of unemployment duration, where the degree of convexity depends on the probability of being unemployed next month. Section 3 applies this assessment protocol to analyse unemployment in Spain for workers from different regions, taking as its reference the data for the beginning of 2015. It provides an estimate of the social cost for different Spanish regions and illustrates well how this assessment protocol improves the vision of unemployment, particularly broadening the problem. A few final comments are given in Section 4 by way of conclusion.

2 The model

2.1. The simplest model

Consider the following extremely simple model of a worker whose utility depends on income and leisure according to the symmetric Cobb-Douglas function given by:

$$u\left(y,\ell\right)=\alpha y^{1/2}\ell^{1/2}$$

where y stands for income, ℓ for leisure, and α is a coefficient that defines the units in which utility is measured. Let T stand for the total amount of time available in a given period to be allocated between labour and leisure and let w denote the corresponding wage rate. Then $y=w\left(T-\ell\right)$, which results in:

$$u\left(y,\ell\right)=\alpha\left[w\left(T-\ell\right)\right]^{1/2}\ell^{1/2}$$

The consumer's optimal choice consists of working for half of the available time and devoting the remaining half to leisure.[3] That is, $y^*=wT/2$, $\ell^*=T/2$, so that $u^*=u\left(y^*,\ell^*\right)=\alpha w^{1/2}(T/2)$. By letting $\alpha=2/T$ the following emerges:

$$u^*=u\left(y^*,\ell^*\right)=w^{1/2}$$

[3] Note that, simple as it is, this model mimics what is a standard behaviour: the 16 hours available in each working day (24 minus 8 devoted to rest) are equally split into 8 hours of work and 8 hours of leisure.

That is, utility in equilibrium can be approximated by the square root of the market wage. This is a simple money metric that becomes the benchmark for the disutility derived from unemployment. When a worker *h* is unemployed he/she may receive unemployment benefit *s* per period for a maximum of *q** periods, provided he/she has earned the pertinent rights. Therefore: [4]

$$u_h^0 = \begin{cases} (s_h)^{1/2} & if\ unemployemnt\ benefit \\ 0 & if\ not \end{cases}$$

The individual utility loss due to unemployment for a worker *h* who "today" has been unemployed for *q* periods and is entitled to unemployment benefits is given by:

$$d_h = \left(u_h^* - u_h^0\right) q_h = \begin{cases} \left((w_h)^{1/2} - (s_h)^{1/2}\right) q_h & if\ q_h \le q^* \\ (w_h)^{1/2} q_h - (s_h)^{1/2} q^* & if\ q_h > q^* \end{cases} \qquad [1a]$$

That is, disutility at a given point in time is measured via an index that reflects the impact on the agent's utility of the corresponding cumulative income loss with respect to being employed. Needless to say this index depends on the units in which wages and unemployment duration are measured.

If the worker has no unemployment benefit then his/her utility loss is given by:

$$d_h = (w_h)^{1/2} q_h \qquad [1b]$$

The following trivial transformation helps to describe the unemployment cost later on in a more general context. Define the cost function c_h as the average income loss of worker *h* when unemployed for q_h periods, that is,

$$c_h(.) = \begin{cases} \left.\begin{matrix} (w_h)^{1/2} - (s_h)^{1/2} & if\ q_h \le q^* \\ \dfrac{(w_h)^{1/2} q_h - (s_h)^{1/2} q^*}{q_h} & if\ q_h > q^* \end{matrix}\right\} with\ unemployment\ benefit \\ (w_h)^{1/2} \quad with\ no\ unemployment\ benefit \end{cases} \qquad [2]$$

where $c_h(.)$ is a shorthanded version of $c(q_h, w_h, s_h)$.

Equations [1a], [1,b] can then be simply rewritten as:

$$d_h = c_h(.) q_h \qquad [1']$$

that is, the average cost per period times the number of periods.

[4] We assume here that each unemployed worker has a constant unemployment benefit for q* periods and nothing afterwards. This is a simplification for the sake of facilitating the exposition. In real life unemployment benefits tend to decrease with duration and there may be some social subsidies for those unemployed for more than q* periods. We return to this point later. The empirical application computes those subsidies and benefits properly.

For a given (active) population N with cardinal n, let U_N denote the set of unemployed in N, with cardinal n^U. The per capita utility loss due to unemployment can be written as:

$$D_N = \frac{1}{n}\sum_{h \in U_N} c_h(.)q_h \qquad [3]$$

where each individual cost function $c_h(.)$ computes all the relevant information regarding agent h (unemployment duration and income loss). Equation [3] can be rewritten in a more intuitive way as:

$$D_N = \frac{n^U}{n} \times \frac{\sum_{h \in U_N} c_h(.)q_h}{n^U} \qquad [3']$$

That is, the per capita social cost of unemployment is an index given by the product of the unemployment rate, n^U / n, and the average disutility of the unemployed person. The first term corresponds to the *incidence* whereas the second one provides a measure of the *severity* of unemployment.[5]

Remark 1: *This formulation implies that getting a job immediately turns unemployment length, and hence the corresponding disutility, into zero. This may appear too rigid an approach for individual data and other interpretations are possible. In particular, the number of months unemployed at time t,* $q_h(t)$*, can be replaced by a function* $\widehat{q}_h(t)$ *of the number of months unemployed in a given time span (e.g. a moving average). Be this as it may, this feature becomes less important when averages are taken over large groups of workers, as is done here for the social assessment. So in principle this question can be dealt with by substituting* $q_h(t)$ *values by suitably chosen* $\widehat{q}_h(t)$ *values.*

2.2.- Convex duration and social subsidies

According to equation [1] the disutility associated with a given unemployment spell is a linear function of its duration. Yet it may be considered that duration should enter disutility as an increasing and convex function, $f\left(q_h\right)$, to reflect the idea that the longer the previous unemployment is, the more a further month of unemployment hurts. The reasons include the cumulative effect of income loss on living standards, the reduction in the probability of finding a new

[5] This expression is very similar to the one proposed in Shorrocks (2009b) for measuring unemployment, for $\alpha = 1$. The key difference is that in our formulation the key reference variable is disutility rather than duration.

job (deterioration of human capital, signalling effect) and increasing difficulties in personal fulfilment and social inclusion.

Assuming that this is the case, the type of convex function *f* that is suitable for this purpose needs to be decided. Recall on this point that the degree of convexity of a function is related to its curvature, which is controlled by its second derivative, usually expressed in terms of the elasticity of the first derivative. In our case that elasticity measures the relative change in the marginal impact of duration on the relative change of individual unemployment length. The simplest constraint to control for the degree of convexity in this context is to assume constant elasticity. This makes it possible to parameterize the impact of unemployment duration by a single number: the value of the elasticity of marginal impact of duration, ν. The function that performs this task is well known and can be expressed as $f\left(q_h\right)=q_h^{1+\nu_h}$, where ν_h stands for the elasticity of the marginal impact of duration for agent *h*.[6] This gives:

$$d_h = c_h(.)q_h^{1+\nu_h} \qquad [4]$$

where $c_h(.)$ is defined as in equation [2]. The convexity of disutility in duration amounts to assuming the "preference for equality" in Shorrocks' framework (i.e. it is better to have two workers unemployed for one month each than one worker unemployed for two months).

Equation [4] describes a family of functions that depend on a single parameter. The next question is how to choose an appropriate value of that parameter. Our proposal here is the following: take ν_h as the probability of agent *h* remaining unemployed for one additional month. That is to say, the degree of convexity of the disutility function is governed by the probability of remaining unemployed. This probability is a measure of the unemployment hysteresis. Needless to say, this is a normative decision that expresses concern for unemployment. Note that this formulation establishes a clear and rather conservative bound on the admissible degree of convexity, as the exponent of the individual disutility function varies between 1 and 2 (i.e. between a linear function and a quadratic one).

The model presented so far adopts a binary description of the income of unemployed persons. That is, they receive unemployment benefits when the duration is below the threshold q^* and nothing from that point onwards. Yet there

[6] This parameter corresponds to the Arrow-Pratt coefficient of relative risk aversion for concave functions (e.g. Pratt, 2013). This is the format adopted by Atkinson (1970) for his reference welfare function, by letting $\varepsilon = -\nu$.

are actually many different situations in real life. In particular, it is often the case that those who remain unemployed and have no right to unemployment benefits, either because they did not contribute enough or because they have exhausted those benefits, have access to social subsidies. Those subsidies usually depend on family needs and can involve a limited or unlimited amount of time.

Let z be the social subsidy and assume that it is incompatible with receiving unemployment benefits and is indefinite, for the sake of simplicity in exposition (the adjustments in the equations are immediately apparent when this is not the case). Equation [2] is transformed into the following:

$$c_h(.) = \begin{cases} \left. \begin{array}{l} (w_h)^{1/2} - (s_h)^{1/2} \quad \text{if } q_h \le q^* \\ \dfrac{(w_h)^{1/2} q_h - (s_h)^{1/2} q^* - (z_h)^{1/2}(q_h - q^*)}{q_h} \quad \text{if } q_h > q^* \end{array} \right\} \textit{with unemployment benefit} \\ (w_h)^{1/2} - (z_h)^{1/2} \quad \textit{with no unemployment benefit} \end{cases}$$

Needless to say, this formulation also includes the case in which there is no social subsidy ($z_h = 0$), which brings us back to equation [2].

The assessment formula in this more general case can thus be expressed by the following index:

$$D_N = \frac{n^U}{n} \times \frac{\sum_{h \in U_N} c_h(.) q_h^{1+v_h}}{n^U} \qquad [5]$$

Now the first term corresponds to the *incidence* and the second provides a measure of the *severity* of unemployment, adjusted for *hysteresis*.

Remark 2: *Note that it is implicitly assumed that n is the size of the active population rather than the number of individuals in society. We believe that this is the right reference. If the whole population, say M with cardinal m, is considered as the reference for calculating the social cost, equation [5] could be rewritten as follows:*

$$D_M = \frac{n}{m} \times \frac{n^U}{n} \times \frac{\sum_{h \in U_N} c_h(.) q_h^{1+v_h}}{n^U}$$

That is, the formula of the social cost would include the participation rate, n/m, as an additional factor.

2.3.- Decomposition

An appealing feature of the assessment formula proposed by Shorrocks (2009b) is that it can be decomposed multiplicatively into the three components that Sen (1976) deems essential in poverty analysis: incidence, intensity and inequality. We show now that our formula can also be decomposed in that way, with one proviso: in our case intensity and inequality refer to disutility and not only to duration (i.e. we also compute the corresponding income loss).

Let $C^U = \frac{\sum_{h \in U_N} c_h(.)}{n^U}$ denote the average income loss of the unemployed, $q_N = \frac{\sum_{h \in U_N} q_h}{n^U}$ the average duration of unemployment in society, and v_N the average probability or remaining unemployed. Now consider the following elementary transformation:

$$\begin{aligned} D_N &= \frac{n^U}{n} \times \frac{\sum_{h \in U_N} c_h(.) q_h^{1+v_h}}{n^U} \\ &= \frac{n^U}{n} \times C^U q_N^{1+v_N} \times \frac{1}{n^U} \sum_{h \in U_N} \frac{c_h(.) q_h^{1+v_h}}{C^U q_N^{1+v_N}} \end{aligned} \qquad [5']$$

The first component of this expression corresponds to the *incidence* of unemployment (the head count ratio, H_N). The second term, $C^U q_N^{1+v_N}$, is a measure of the *intensity* of unemployment, S_N, given by the average disutility of the unemployed. Finally, the third term is a measure of *inequality* in disutility among the unemployed, I_N, which is given by the sum of the shares of individual disutility in average disutility. To get an inequality measure that yields a zero value when all disutilities are equal, take the following:

$$I_N = \frac{1}{n^U} \sum_{h \in U_N} \frac{c_h(.) q_h^{1+v_h}}{C^U q_N^{1+v_N}} - 1$$

so that our assessment formula can be decomposed as follows:

$$D_N = H_N \times S_N \times (1 + I_N) \qquad [6]$$

Equation [6] is the precise counterpart of Shorrocks' decomposition. It says that the social cost of unemployment can be interpreted as the product of the incidence and intensity of unemployment, with a penalty factor that reflects the inequality in disutility among the unemployed.

3 Implementation: the case of Spain

We now apply this assessment protocol to the Spanish labour market at the beginning of 2015, focusing on the differences between the seventeen autonomous regions. Our empirical work relies on the use of two different databases: one for employed workers and the other for the unemployed. In the first case we work with a representative sample of about 170,000 observations whereas for the second we use the whole census of unemployed workers, with more than five million observations.

3.1 Data

The dataset for employed workers is the *Spanish Earnings Structure Survey* (SESS), which contains detailed micro-data on the characteristics of employed workers and the various components of their wages. The *Earnings Structure Survey* (ESS) is a European individual dataset with harmonised information for a bundle of European countries. The waves available are 2006, 2010 and 2014. The information available on demographic characteristics includes gender, age, educational attainment and whether each worker is foreign or native. Labour market characteristics contain information about the type of contract, tenure in the firm, occupation and sector of activity, hours worked (including overtime) and detailed information on wages, such as the base wage, overtime pay and other complements. The sample consists of 169,062 full-time workers in the 2014 wave and the survey includes a weighting factor that enables the sample to be weighted for population inference purposes. Our focus is to obtain estimates of the gross hourly wages for the different types of worker. The range of the hourly wage was set between 2 and 60 euros.[7]

Our second dataset consists of monthly longitudinal information on all individuals registered with the *Spanish Public Employment Service* (SPES) from January 2011 to September 2017. Data are collected on the last day of each month. Most of those registered are unemployed, but some may be employed and searching for another job (their employment status is clearly stated, though). The database includes all the information provided by each individual when registering at the employment office, including standard demographic characteristics (gender, age, education level, nationality, postcode and residence, knowledge of other

[7] Those workers earning less than €2/hour account for 0.76% of the sample and those earning more than €60/hour for 0.91%.

languages), along with labour market information (previous employment experience, occupational and geographical searches, unemployment duration, etc.). The SPES also provides precise information on the type of unemployment benefits or social subsidies that individuals are receiving or last received and the start and end dates of their entitlement. There is no information on the amount received as benefits, but this can be inferred as unemployment benefits correspond to a (time varying) proportion of wages and social subsidies are a proportion of the of Multiple Effects Public Income Indicator.

Our dataset contains all individuals registered as looking for a job in January 2015 (5,520,253 persons).

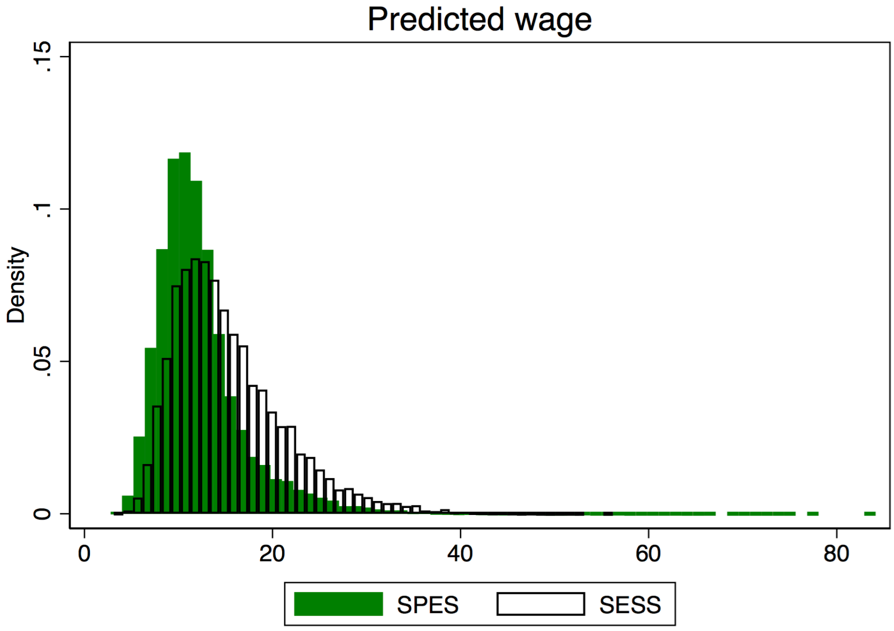
The first step towards approaching the social cost of unemployment is to estimate the loss of wages. To that end we consider those variables contained in both datasets which we know are important determinants for wages, i.e.: gender (2 groups), age (10 groups), level of education (10 groups), a dummy indicating whether the individual is foreign or native, sector of activity (19 groups) and 2-digit sector of occupation (58 groups). Using the SESS, we estimate hourly wages and obtain the predicted hourly wage for every worker in the SESS sample. Then we impute that predicted wage to all workers registered as unemployed in January 2015 in the SPES, on the basis of their gender, age, level of education, nationality, former sector of activity and former occupation. [8] To be more precise, we create cells (2x10x10x2x19x58=440800 cells) from the categories defined for gender, age, education nationality, sector and occupation, and assign an imputed wage for each cell based on the above wage prediction[9]. As a result, two unemployed workers in January 2015 belonging to the same cell would have the same imputed wage.

The distribution of the predicted wages for the 2014 SESS workers and for the unemployed individuals is presented in Figure 1, where the differences in the shapes correspond to the different compositions of the two groups.

[8] We drop unemployed individuals with no previous employment experience given that their wages cannot be imputed in the same way and as a group they may have very different characteristics. They account for 0.38% of unemployed individuals. We also drop those unemployed individuals who only seek part-time work, as their disutility function might be different. They account for 0.94%.

[9] Following the recommendation of López-Laborda, Marín-González and Onrubia (2017), we use a Generalized Linear Model to estimate the predicted wage in order to avoid bias in the estimation results due to the retransformation problem from logarithms to wage levels.

Figure 1. Distribution of predicted hourly wages in the 2014 SESS and for the unemployed in the SPES, January 2015.



In line with to the hourly wage imputed, we estimate the monthly wage as 22 (days/month) x 8 (hours/day) x hourly wage (€/hour). From the monthly individual information on types of unemployment benefit and unemployment duration, we impute the amount of unemployment benefit that each unemployed individual is receiving and compute the average cost of unemployment for each unemployed worker. More precisely, the monthly unemployment benefit is calculated as 70% of the monthly wage for the first 180 days and 50% of the monthly wage for the following months in which it is received. It is upper and lower bounded at €1411.83 and €501.98, respectively. The amount corresponding to social subsidies is 75%, 80% or 107% of the Multiple Effects Public Income Indicator (set at €532,51)[10] depending on the type.

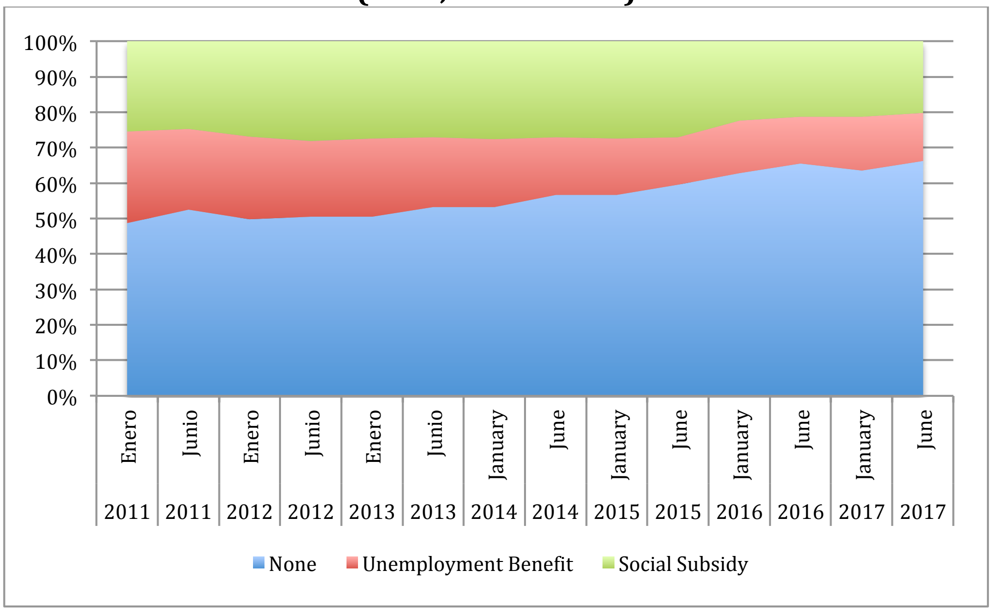
Next we estimate the probability of individuals finding a job in the next month (by the last day of one month for all those unemployed on the last day of the previous month). To that end we use a discrete choice model where the dependent variable takes a value of 1 if individuals find work in the next month and zero if

[10] The upper and lower bounds and the social subsidies depend on the Multiple Effects Public Income Indicator, which has remained unchanged at €532.51since 2011.

they remain unemployed. We estimate a probit model to calculate the probability of a job being found for every month from January 2015 to February 2015. To that end we take into account all observable variables that may affect the employability of those registered with the Public Employment Service. In particular, we include demographic characteristics such as gender, age, nationality, disability, education and language skills; job characteristics such as unemployment duration, occupations requested, experience, activity in the previous field of work, geographical scope of the new job search and region of registration.

After estimating the imputed wage and the probability of remaining unemployed for one more month depending on the characteristics of each unemployed individual (4,988,632 agents in our dataset in January 2015[11]), we consider three different groups of unemployed workers: (1) those who receive unemployment benefits (UB); (2) those who receive social subsidies (SS); and (3) those who receive no income (N). Figure 2 shows the trends in these three groups of unemployed persons from 2011 to 2017.

Figure 2. Distribution of the unemployed depending on income sources (SPES, 2011-2017).



The disutility of an unemployed worker h who receives unemployment

[11] We drop from our database all those unemployed agents older than 64 as they are considered as inactive, those who became unemployed in January 2015 (whose unemployment duration is zero) and those who are not observed in February 2015 and whose probability of remaining unemployed cannot be determined. As explained above, we also disregard all unemployed persons with no prior experience and those looking for part-time jobs. Finally, we also drop those individuals who became unemployed before 2011 whose unemployment benefits before 2011 cannot be imputed (they account for 1.66%).

benefits is obtained by directly applying the corresponding formula, $\left(w_h - s_h\right) q_h^{1+v_h}$. The richness of the dataset enables the monthly disutility to be computed for each unemployed individual since their entry into unemployment, according to the type of unemployment benefit that they are receiving.

Among the group of unemployed workers who have received social subsidies at any time, three different situations can be found: (a) unemployed workers who have exhausted their unemployment benefits and then receive a social subsidy; (b) unemployed workers who have been receiving a social subsidy throughout their period of unemployment; and (c) unemployed workers who started receiving social subsidies after a period of not receiving any benefit in 2015.

Similarly, those receiving no payments fall into four types: those who have exhausted unemployment benefits, those who have received social subsidies for a period and ceased to receive them, those who received unemployment benefit, then social subsidies but have exhausted both and those who have never received any payments.

3.2 Empirical Results: Computing The Social Costs of Unemployment for Spain's regions

Now we present the main results on the Spanish labour market, using the 2014 data on wages (last available wave from the *Spanish Earnings Structure Survey*) and those of January 2015 for the Spanish Register of Unemployed Workers. The empirical analysis refers to a single period and focuses on comparing (per capita) social costs in Spain's regions. Recall that Spain is a highly decentralised country where control of about half of public expenditure is devolved to the regions. Furthermore, many areas of public authority such as health, education and other economic activities are also devolved to the regions. It is therefore interesting to analyse how the social cost of unemployment varies from one region to another and how the relative cost differs from the relative unemployment rates[12]. This comparison clearly illustrates that focusing on unemployment rates may give a distorted view of what unemployment implies for society. Indeed, the coefficient of correlation between unemployment rates and

[12] One of the richnesses of the model presented is that it can also be used to analyse how social cost varies across different types of worker.

social costs is negative (about -0.3), which already gives a first hint as to the differences in behaviour of the two variables.

The first five columns of Table 1 provide the key data on the regions: population shares, unemployment rates and average figures for duration, cost and probability of remaining unemployed. It must be remarked here that for the sake of consistency in the use of the database, the term "unemployment rate" is used for the ratio of the number of persons registered as unemployed to the active population.[13] The last two columns contain the assessment of the social cost of unemployment, as described by equation (5), in both absolute values (under the heading "Value") and relative terms by normalising the values and setting the Spanish average to 100. Recall that in our model "per capita" means per active worker (see Remark 2 above).

Table 1: Duration, income loss, probability of remaining unemployed and social cost of unemployment by regions (Spain, 2015)

	Population share	Unemployment Rate	q	c(.)	v	Social cost (per capita)	
						Value	Relative value
Total	100%	18.17%	18.83	35.32	0.964	5,526	100
Andalusia	21.3%	21.92%	24.52	36.70	0.961	9,257	168
Aragón	2.3%	14.85%	15.81	34.04	0.963	3,410	62
Asturias	2.3%	20.85%	19.93	37.08	0.974	7,800	141
Balearic Islands	3.0%	21.27%	8.60	30.59	0.936	1,849	33
Canary Islands	5.4%	20.15%	18.21	33.40	0.968	5,094	92
Cantabria	1.2%	17.96%	16.39	34.84	0.922	4,407	80
Castilla Mancha	4.6%	16.68%	18.00	35.95	0.971	5,252	95
Castilla León	5.5%	23.08%	15.67	33.98	0.969	4,567	83
Catalonia	13.6%	14.96%	16.55	32.66	0.966	3,596	65
C. Valenciana	11.6%	19.90%	17.43	34.58	0.969	4,774	86
Extremadura	3.1%	25.70%	14.86	34.66	0.970	4,682	85
Galicia	5.6%	18.44%	18.54	35.05	0.966	6,197	112
Madrid	11.5%	14.11%	17.38	35.52	0.969	4,008	73
Murcia	3.2%	19.01%	15.93	34.16	0.970	3,902	71
Navarre	1.1%	14.95%	15.64	33.56	0.961	3,340	60
Basque Country	3.5%	13.98%	28.37	40.55	0.950	10,456	189
Rioja	0.6%	16.22%	15.38	33.21	0.965	3,433	62
Coefficient of variation		0.181	0.232	0.060	0.014	0,428	

[13] These figures are often lower than the conventional unemployment rates measured via the standard survey (Encuesta de Población Activa, in Spain). The difference for the whole country is some five points, up from 18.17% to 23.78%, and the coefficient of correlation between the two series is 0.8. Note, though, that this change does not affect the estimates of social cost, as the unemployment rate is introduced as an artefact to make the assessment formula easier to interpret.

The last row shows that there are large disparities between regions in both unemployment rates (ranging from 26% to 14%) and duration (ranging from 28.4 to 8.6). These disparities accumulate when the social cost of unemployment is considered. Andalusia, the Balearic Islands, Castilla y León and Extremadura are the regions with the highest unemployment rates, but they differ substantially in average duration (24.5 for Andalusia and 8.6 for the Balearics). The relative social cost clearly illustrates those differences: it ranges from 168% of the total for Andalusia to 33% for the Balearics. Also observe that Aragón, the Balearics, Catalonia, Madrid, Murcia, Navarre and Rioja all show social costs below 75% of the total and yet their unemployment rates differ substantially (from 21.3% in the case of the Balearics to 14.1% in the case of Madrid). In summary, unemployment rates are very poor proxies of the impact of unemployment within regions.

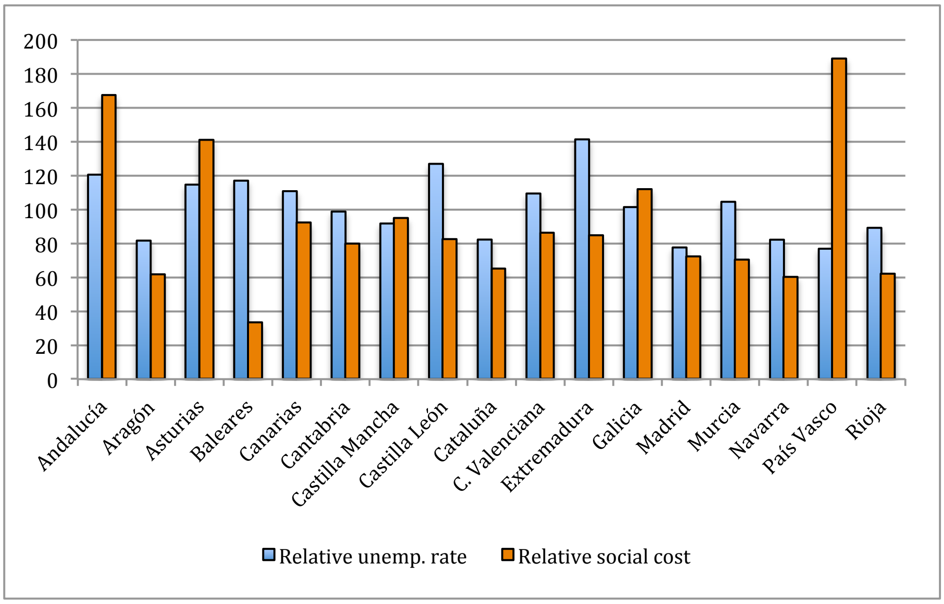
We find similar results in the regions with the best figures. Unemployment rates in Madrid and the Basque Country are less than 80% of the mean. Yet there is much greater duration in the Basque Country. As a result, when these data are combined with the corresponding income loss, the relative social cost for those regions ranges from 189% for the Basque Country Vasco and 73% for Madrid. Here again unemployment rates are found to hide the impact of unemployment on social welfare.

The probability of remaining unemployed is very high in all regions (over 0.92) with very slight variability[14]. Average costs also show a rather homogeneous behaviour across the regions.

Figure 3 illustrates well how different unemployment is perceived to be when unemployment rates are considered from when the corresponding social costs are computed. The graphic compares the relative values of the two variables for the regions, setting the Spanish average at 100. It can be seen that Andalusia, Asturias and most prominently the Basque Country have costs which are relatively much higher than their corresponding unemployment rates. The contrary is the case in Aragón, the Balearic Islands, the Canary Islands, Castilla León, Comunidad Valenciana, Extremadura and Murcia.

[14] It is worth mentioning that the choice of coefficient 2 in Goerlich & Miñano (2018) is practically the same as that obtained from using the probability of remaining unemployed.

Figure 3: Relative unemployment rates and relative social costs for Spain's Regions (2015)



3.3 Decomposition

We now present the decomposition of the social cost of unemployment into three components – incidence, intensity and inequality – according to equation [6]. Table 2 provides the basic data in both absolute and relative terms. The first point to note is, once more, the broad diversity between Spain's regions. The biggest gaps are in intensity (i.e. duration), with a figure more than double that for inequality, which in turn is 1.4 times greater than that for incidence. Incidence is negatively correlated with both intensity (-0.19) and with inequality (-0.17); inequality and intensity show a stronger negative correlation (-0.5). Andalusia, Asturias and most notably the Basque Country show the highest figures for intensity, well above the Spanish average. On the opposite side are Aragón, the Balearics, Cantabria, Castilla León, Extremadura, Navarre and Rioja, at more than 30 points below the mean. For inequality, the Balearics, Cantabria, Catalonia, Galicia, Navarre and Rioja are the regions with the highest figures (more than 15 points above the mean), while Andalusia, the Canary Islands and Comunidad Valenciana have the lowest figures.

Table 2: Decomposition of social cost

	Incidence H_N	Intensity S_N	Inequality I_N	Relative incidence	Relative intensity	Relative inequality
Spain	18%	11,267	1.70	100	100	100
Andalusia	22%	19,477	1.17	122	173	69
Aragón	15%	7,682	1.99	82	68	117
Asturias	21%	13,626	1.75	116	121	103
Balearic Islands	21%	1,971	3.41	118	17	200
Canary Islands	20%	10,093	1.51	112	90	89
Cantabria	18%	7,525	2.26	100	67	133
Castilla Mancha	17%	10,711	1.94	93	95	114
Castilla León	23%	7,662	1.58	128	68	93
Catalonia	15%	8,132	1.96	83	72	115
C. Valenciana	20%	9,615	1.50	111	85	88
Extremadura	26%	7,058	1.58	143	63	93
Galicia	18%	10,909	2.08	102	97	122
Madrid	14%	9,820	1.89	78	87	111
Murcia	19%	7,978	1.57	106	71	93
Navarre	15%	7,374	2.03	83	65	119
Basque Country	14%	27,610	1.71	78	245	101
Rioja	16%	7,139	1.97	90	63	116
C V	0.181	0.542	0.248			

3.4 Sensitivity analysis

The formula used to assess the social cost of unemployment involves two particular assumptions on the disutility of the unemployed: (a) disutility is a convex function of duration, with a degree of convexity associated with the probability of remaining unemployed, $q_h^{1+v_h}$. (b) Utility is concave in income: $u_h* = w_h^{1/2}$. Even though both are conventional assumptions it is interesting to analyse their effects on the empirical results. Hence we now provide new estimates of the social cost of unemployment in three alternative scenarios: first, assuming that disutility is linear in duration (i.e. using equation [3'] to assess social cost); second, assuming that utility is linear in income (which amounts to taking the monthly wage, rather than its square root, as the proper money metric for utility); and third, assuming that utility is linear in both duration and income. Clearly, relative differences will be positive or negative depending on the impact of removing hysteresis and/or linearity in income in each region relative to the change in the whole country.

Table 3 provides the results in relative terms (i.e. setting the value for Spain at 100). The first column replicates the last one in Table 1, to facilitate the comparison of results. The second column is obtained by ignoring the hysteresis factor (i.e. when it is assumed that $v_h = 0 \ \forall h$). The third column reintroduces hysteresis but makes utility linear in income (i.e. $u_h{}^* = w_h$). The last column shows the case in which both changes are introduced, i.e. no hysteresis and utility linear in income.

Table 3: Relative social cost of unemployment by regions (Comparison between alternative scenarios. Spain, 2015)

	Reference model	No hysteresis	Linear income	No hysteresis and linear income
Total	100	100	100	100
Andalusia	168	163	159	154
Aragón	62	66	64	68
Asturias	141	128	146	132
Balearic Islands	33	46	35	48
Canary Islands	92	101	92	100
Cantabria	80	85	80	85
Castilla Mancha	95	89	96	91
Castilla León	83	102	83	101
Catalonia	65	67	69	71
C. Valenciana	86	99	88	100
Extremadura	85	110	85	108
Galicia	112	99	113	102
Madrid	73	72	77	77
Murcia	71	86	70	83
Navarre	60	65	62	67
Basque Country	189	133	186	131
Rioja	62	69	63	70
CV	0.428	0.297	0.396	0.276

The data show that making utility a convex function of duration has a relevant impact on the assessment. The reason is that there are large differences in duration whereas the probability of remaining unemployed is close to 1 and highly homogeneous. The Balearic, Castilla León, Extremadura, Murcia and the Basque Country show differences of more than 15 points when disutility is made linear in duration (with a positive effect for the Basque Country and a negative one for the other four regions). Making utility linear in income has a very small impact on the assessment (all relative changes are below 5%). This is reflected clearly in the coefficient of variation (last row of Table 3).

4 Final remarks

We present a protocol for assessing the social cost of unemployment that involves three different dimensions – incidence, severity and hysteresis – integrated into a single-value indicator. That indicator is obtained as a social welfare function that aggregates individual disutilities, which depends on unemployment duration, income losses and the probability of remaining unemployed. The assessment formula thus obtained is simple and intuitive and can be expressed as the product of two different variables: the conventional unemployment rate (a measure of incidence) and the average cost of unemployment (a measure of severity adjusted for hysteresis). Alternatively, the assessment can be decomposed as with conventional poverty measures into incidence, intensity and inequality, in line with the work of Shorrocks (2009b).

We have used this protocol to analyse unemployment in Spain at a given point in time (January 2015), comparing the situation of the country's various regions. Note that collecting the data required to compute the social cost of unemployment is already an interesting exercise that provides a particular angle from which to approach the problem.

Our empirical analysis shows that this way of approaching the measurement of unemployment provides new insights into the nature and extent of the problem. In particular, it can be seen that regional disparities in duration are much greater than those in unemployment rates, which translates into large differences in the social cost of unemployment that go in different directions depending on the region. As illustrated in Figure 3, a very different picture of the unemployment problem emerges depending on whether unemployment rates or the corresponding social costs are considered. The driving force behind those differences is duration, which is amplified when it is assumed that the longer the unemployment spell lasts the more a further month of unemployment hurts. Summarising, focusing only on unemployment rates might give a distorted view of what unemployment implies for society. The model presented helps to provide a deeper understanding of the social cost of unemployment for different population groups.

References

Atkinson, A. B. (1970). On the measurement of inequality. *Journal of economic theory*, *2*(3), 244-263.

Bentolila, S., García-Pérez, J. I., & Jansen, M. (2017). Are the Spanish long-term unemployed unemployable?. *SERIEs*, *8*(1), 1-41.

Chakravarty, S. R. (2009). Inequality, polarization and poverty. *Advances in distributional analysis. New York*.

Dalton, H. (1920). The measurement of the inequality of incomes. *The Economic Journal*, *30*(119), 348-361.

de La Rica, S., & Gorjón, L. (2017). Assessing the Impact of a Minimum Income Scheme in the Basque Country, mimeo.

Goerlich, F.J. & Miñano, A. (2018), Unemployment, spell duration and the burden of unemployment in Spain during the XXI Century, University of Valencia, mimeo.

Jones, C. I., & Klenow, P. J. (2016). Beyond GDP? Welfare across countries and time. *American Economic Review*, *106*(9), 2426-57.

López-Labordaa, J., Marín-Gonzálezb, C., & Onrubiac, J. (2017). Estimating Engel curves: A new way to improve the SILC-HBS matching process. *Documento de Trabajo, 15*.

Pratt, J. W. (2013). Risk aversion in the small and in the large. In *HANDBOOK OF THE FUNDAMENTALS OF FINANCIAL DECISION MAKING: Part I* (pp. 317-331).

Sen, A., & Foster, J. E. (1973). *On economic inequality*. Oxford University Press.

Sen, A. (1976). Poverty: an ordinal approach to measurement. *Econometrica: Journal of the Econometric Society*, 219-231.

Sengupta, M. (2009). Unemployment duration and the measurement of unemployment. *The Journal of Economic Inequality*, *7*(3), 273.

Shorrocks, A. (2009). On the measurement of unemployment. *The Journal of Economic Inequality*, *7*(3), 311-327.

Shorrocks, A. (2009). Spell incidence, spell duration and the measurement of unemployment. *The Journal of Economic Inequality*, *7*(3), 295-310.

Villar, A. (2017). *Lectures on inequality, poverty and welfare*. Springer International Publishing.