An econometric investigation into the nature of the relationship between unemployment and suicide

Bulat Kubeyev
Senior Sophister

With this comprehensive econometric study, Bulat Kubeyev investigates the effect of unemployment levels on suicide rates at different levels of GDP/capita, using data from 88 countries over the last decade. Interestingly, it is found that a country's income per person has a significant effect on the direction of the relationship between suicide rates and unemployment.

Introduction

'You take my life, when you do take the means whereby I live.'
      -The Merchant of Venice (4.1.371-72)

Dorland's Illustrated Medical Dictionary defines suicide as 'the act of intentional taking of one's own life'. It can be triggered by a wide range of social and economic factors, and while the study of suicide sometimes presents a challenge to evaluate different aspects of societal behavior (e.g. bullying, loneliness, etc.), suicide rates, like most economic situations, can be proxied by unemployment rates. The effects of unemployment on suicide have undergone some major quantitative analyses? for Sweden, Germany and the United States, while mostly qualitative analysis has been performed for Ireland.

      But does unemployment really affect suicide rates? Previous research in this area has not provided a definite answer, but most macroeconomists would agree that unemployment results in a loss of social and financial stability. This generally leads to severe frustration or even depression, because 'unemployed people not only lose materially, they also potentially lose access to social networks, self-esteem, self-confidence, a scheduled life structure, a sense of identity and possibly a purpose for their lives' (Neumayer, 2004,

70

p.1038). Naturally, depressed persons are more prone to suicide than those with no underlying mental health problems.

A factor that was previously overlooked in most research papers is the link between unemployment and income. However, though the effect of unemployment on suicide rates seems to be intuitively quite straightforward, when it is interacted with personal income the results become somewhat ambiguous. This paper is examining whether the effect that unemployment has on suicide rates varies consistently with the level of real GDP per capita.

It is worth mentioning at this stage that the study presented in this paper is not the first of its kind and hopefully not the last. Still, one of its unique features is its geographical scale, while its main contribution is that it allows us to understand the effects of unemployment on suicide in countries with various income levels, which are not necessarily similar[1].

Background/ Motivation

The idea that there is an association between unemployment and suicide is hardly new. Much literature has been written on the subject and it can certainly be grouped by methods of analysis, countries of interest and even by the hypothesis being tested. Nevertheless, practically all significant research in this area has one common feature – it was conducted using either time-series or cross-sectional data analysis. This research proves to be invaluable as it is generally thorough and definitive, but it is mostly conflicting. For instance, Gerdtham and Johannesson (2003) use time series analysis for Sweden and end up with the positive relationship between unemployment and suicide rates. However, when Neumayer (2004) applies the same method of time series analysis to Germany, this relationship is negative[2].

Generally speaking, most recent papers, such as the one by Andres (2005), use ordinary least squares (OLS) as a method of analysis and then go on to report the positive and significant effect that unemployment has on aggregate suicide rates, but in the end fail to find any significant association between income and suicide. On the other hand, Chuang and Huang (1997) find this relationship to be negative, if only across Taiwan.

Similarly, at the initial stages of research, it was assumed that this

1 The idea of "similar income levels" is abstract in this context: EU and OECD member states all have different income levels, but they are also all developed and can be thought of as being "similar". Hence, including third world countries and CIS member states can improve the transparency of the results.

2 It may be argued that the discrepancy in the results of similar analyses for Sweden and Germany is caused by the differences between The Nordic Model and The Rhineland Model respectively, but such an argument would be mere speculation. Other factors (e.g. amount of sun light per year) affect one's mental health just as much.

paper would focus solely on Ireland (including the Republic of Ireland and Northern Ireland), partially due to the scope of the information that was available in the public domain. A national study conducted in 2001 on behalf of the Irish Health Board provides a fully qualitative assessment of suicide trends in Ireland complete with the profiles and social standings of the deceased. However, with a population of approximately 6.2 million people and only 32 counties, it soon became clear that Ireland was a poor candidate for this type of study. Furthermore, even if a time-series analysis were to be conducted, collecting annual figures related to suicide (complete with the profiles of the deceased) would present a significant challenge, as the Republic's national police force, an Garda Síochána, can only disclose details with the permission of the families of the deceased.

It immediately becomes evident from the existing research that once the effects of unemployment or income have been estimated, they largely depend on both the country and the time period in question. Furthermore, one can expect a certain degree of interaction between these variables, which results in biased estimators. According to Wooldridge (2003), ignoring a significant interaction term usually produces higher variances and thus, reduces efficiency of the regression model. One possible solution to this problem is to introduce an interaction term between two main control variables and see what effect it has on suicide rates, i.e. to include two effects at the same time. It makes sense to further control for certain demographic and social factors, which might influence aggregate suicide rates.

In general, this paper will try to estimate a fixed effects model, by way of panel-data analysis, for 88 countries with different development levels (including member states of the Commonwealth of Nations, EU, OECD, UN-ASUR, CIS and EFTA) during ten consecutive years (2000-2009). The decision to use panel-data analysis, instead of a cross-section or a time-series, was made with a view to controlling for potential bias associated with omitting a variable or failing to include a time invariant factor that is unique to a particular country (e.g. low levels of yearly sunlight in northern countries quite possibly affects suicide rates and yet remains unobserved).

Empirical Approach

The objective of this paper is to find out how unemployment affects suicide rates depending on income. To do so, it analyzes the relationship between suicide rates and a number of social and economic factors in 88 countries during 2000-2009. Consequently, it compensates for the problem of omitted variable bias.

However, before constructing an empirical model, it is vital to make a decision as to which variables should be included and which should be omitted altogether from the final equation. The decision process is mostly concerned with the independent variables, also known as control variables, since the only possible choice for the dependent variable is the aggregate suicide rate, "$srate_{i,t}$".

As regards the explanatory variables, it is logical to include unemployment rates "$unemp_{i,t}$" and real GDP per capita "$gdppc_{i,t}$", so that we are able to test our hypothesis. Furthermore, as suggested by Wooldridge (2003), an interaction term "$interact_{i,t}$" is introduced in the equation to examine whether the relationship between unemployment and suicide varies for countries with different income levels. This term is generated using STATA 11 and thus, is not automatically present in the dataset.

Now, even though unemployment rates are themselves sufficient to control for the fluctuations in the economy, it is advised to include another variable that serves the same purpose. According to Andres (2005), this variable is the real GDP growth rate, "$grgdp_{i,t}$".

The remaining three variables are introduced in order to control for social factors. These variables tend to change over time and are thought to have a direct impact on the aggregate suicide rates. Using the idea put forward by Chuang and Huang (1997) with reference to Durkheim (1897) that societal suicide rates are influenced by social integration and social regulation, both fertility rate "$fertility_{i,t}$" and family formation rate "$ffrate_{i,t}$" have been included for each country. Finally, alcohol consumption has been previously used as a control variable by Andres (2005). Nevertheless, some experts remain sceptical on the subject of alcohol's relationship to depression, and consequently suicide. The decision to include "$alcohol_{i,t}$" was based on the idea put forward by Chaloupka, Grossman and Saffer (2002), that alcohol consumption causes depression and not vice versa.

Since the study presented in this paper is conducted using panel-data analysis, it is vital to include a full set of year dummies (and to leave one out to compensate for the intercept) in order to control for aggregate time effects. Thus, the baseline equation to be estimated is:

$$srate_{i,t} = \beta_0 + \beta_1 unemp_{i,t} + \beta_2 gdppc_{i,t} + \beta_{IN} interact_{i,t} + Y_{i,t}\lambda + \alpha_i + \gamma_t + \varepsilon_{i,t}$$

(1)

Where $Y_{i,t}$ is a vector, which denotes the following control variables: "$grgdp_{i,t}$", "$fertility_{i,t}$", "$ffrate_{i,t}$" and "$alcohol_{i,t}$". Subscripts i and t in equation (1) index

a country and a time period respectively; β_0, β_1, β_2, β_{IN} are estimated coefficients; λ denotes an estimated coefficient vector; α_i and γ_t are dummy variables whose purpose is to capture unobserved country and time specific effects respectively; $\varepsilon_{i,t}$ is an error term (and we further assume that $\varepsilon_{i,t}$ is independently and identically distributed with mean 0 and variance σ^2 for all i and t).

It is worth noting that if the interaction term is significant, then the partial effect of "$unemp_{i,t}$" on "$srate_{i,t}$" will depend on the level of "$gdppc_{i,t}$" (per Wooldridge, 2003:190):

$$\Delta E(srate_{i,t})\Delta unemp_{i,t} = \beta_1 + \beta_{IN}gdppc_{i,t} \qquad (2)$$

Consequently, all else being equal, the marginal effect of unemployment varies with different levels of real GDP per capita. Similarly, the marginal effect of real GDP per capita changes with various unemployment rates when everything else is held constant:

$$\Delta E(srate_{i,t})\Delta gdppc_{i,t} = \beta_2 + \beta_{IN}unemp_{i,t} \qquad (3)$$

The primary focus is on the relationship described by equation (2).

Dataset

The dataset, comprised of annual suicide rates and socio-economic figures for 88 countries during 2000-2009, has been compiled from the following three sources: WHO Mental Health Data 2011, World Bank Database 2011 and the United Nations Population Division 2011.

Although this dataset is quite large geographically, it has been carefully selected. It may be argued that by including countries with different development levels, one compromises both data transparency and quality of indicators. However, all 88 countries in the dataset have their own well-established national statistics agencies, which have been consistently supplying quality data to their international counterparts. Countries with unique social features have been excluded in order to achieve "truer" coefficients (e.g. Maghreb states consume very little alcohol compared to the rest of the world – they were omitted).

All descriptive statistics for the variables used in equation (1) are listed in Table 1, along with their respective meanings (generated using STA-TA 11).
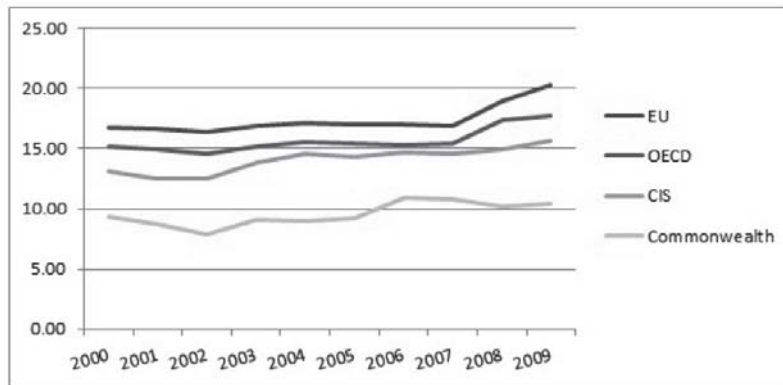
Table 1: Variables and general statistics

Variable	Meaning	Observ.	Mean	St. Dev.
srate$_{i,t}$	Suicide rate (per 100,000)	880	11.07	8.89
unemp$_{i,t}$	Unemployment, total (% of total labour force)	880	9.28	6.20
gdppc$_{i,t}$	GDP per capita (current US $)	880	13924.37	16893.88
grgdp$_{i,t}$	Growth rate of GDP per capita (%)	880	3.22	4.65
alcohol$_{i,t}$	Adult consumption of alcohol, per capita (litres)	880	7.06	3.85
ffrate$_{i,t}$	Family formation rate (per 1000)	880	5.51	1.82
fertility$_{i,t}$	Fertility rate, total (births per woman)	880	2.05	0.72

Sources: WHO 2011; World Bank 2011; UN 2011.

General trends associated with suicide rates in four major economic and political entities are presented in Figure 1.

Figure 1: Suicide rates (per 100,000) during 2000-2009

Dataset Issues

In his paper, Andres (2005) states:

> "The effect of socioeconomic predictors of suicide differs across age groups. Thus, the use of age-specific suicide data definitely does make sense, and as demonstrated in this study, it may result in helpful guidance for health policy makers"
>
> (Andres, 2005, p.449)

Due to the lack of age- and gender-specific statistics, the dataset used here is not as comprehensive as was initially intended. However, it is not the objective of this paper to investigate the effects of unemployment on suicide rates for different age groups and genders. On the contrary, it aims to study the aggregate effect of unemployment on suicide rates with respect to different income levels. Therefore, omitting age-specific statistics does not constitute a major drawback in this type of research.

Empirical Results

$$srate_{i,t} = 17.032 - 0.331unemp_{i,t} + 0.0007gdppc_{i,t} + 0.0008interact_{i,t}$$
$$(1.926) \quad (0.055) \quad\quad (0.00002) \quad\quad (0.0004)$$

$$- 0.187grgdp_{i,t} + 0.399alcohol_{i,t} - 0.472\,ffrate_{i,t} - 2.435fertility_{i,t}$$
$$(0.026) \quad\quad (0.089) \quad\quad (0.181) \quad\quad (0.807)$$

$$N = 880 \quad R\text{-sq} = 0.5455 \text{ (within)}$$

The intercept term is 17.032, which represents the suicide rate if all independent variables were zero simultaneously – not a meaningful situation in itself. The directions of the relationships, indicated by the signs of the coefficients of the independent variables and the intercept, are largely as anticipated. Specifically, all else being equal, a one percent increase in alcohol consumption corresponds to a 0.399 percent increase in suicide rate. On the other hand, a one percent increase in family formation rate corresponds to a 0.472 percent decrease in suicide rate, ceteris paribus. Similarly, holding all else constant, a one percent increase in fertility rate results in a 2.435 percent drop in suicide rate. The coefficient on economic growth, although significant by intuition, is statistically insignificant with a p-value of 0.346. This is an odd occurrence, considering that it was previously used in the study of suicide by other re-

searchers.

Most importantly, in the presence of both "$unemp_{i,t}$" and "$interact_{i,t}$", the negative impact of unemployment on suicide is reduced by income, since $\beta 1 < 0$ while $\beta IN > 0$. On the other hand, in the presence of "$gdppc_{i,t}$" and "$interact_{i,t}$", the effect of income on suicide rates does not change. The beta value for the interaction term is statistically different from zero at the 1% significance level. Hence, the partial effect of "$unemp_{i,t}$" on "$srate_{i,t}$" depends on the level of "$gdppc_{i,t}$". Consequently, it was decided to run the original regression again, this time including a dummy variable that takes on the value of one for the countries with high levels of real GDP per capita and zero otherwise. Note that "$grgdp_{i,t}$" was omitted. The following regression was obtained:

$$srate_{i,t} = 16.534 - 0.364 unemp_{i,t} + 0.0006 gdppc_{i,t} + 0.001 interact_{i,t}$$
$$(1.973) \quad (0.058) \quad (0.00002) \quad (0.00038)$$

$$+ 0.386 alcohol_{i,t} - 0.49\, ffrate_{i,t} - 2.341 fertility_{i,t}$$
$$(0.09) \quad (0.182) \quad (0.811)$$

$$N = 880 \quad R\text{-}sq = 0.6472 \text{ (within)}$$

After removing the interaction term, the resulting regression is:

$$srate_{i,t} = 16.073 + 0.358 unemp_{i,t} + 0.0004 gdppc_{i,t} + 0.274 alcohol_{i,t}$$
$$(1.964) \quad (0.053) \quad (0.00001) \quad (0.0901)$$

$$-0.470 ffrate_{i,t} - 2.369 fertility_{i,t}$$
$$(0.182) \quad (0.812)$$

$$N = 880 \quad R\text{-}sq = 0.6420 \text{ (within)}$$

Therefore, in richer countries an increase of one percent in unemployment rate corresponds to a ceteris paribus 0.36 percent rise in suicide rate.

The R-squared value for the regression with a dummy variable and an omitted variable is greater than its value for the original regression. Arguably, this indicates that the latter model has better predictive powers than the former.

However, the most important result obtained was that for countries with higher income levels the effect of unemployment on suicide was indeed positive and significant. Conversely, for countries with lower incomes, this

effect is negative.[3]

Possible Extensions

The results of the empirical analysis seem to be conclusive. Nevertheless, a further study of the effect that each control variable has on suicide rates is certainly possible. Naturally, in order to take full advantage of the panel-data analysis, the proposed study should cover a much longer time period. Furthermore, it would be interesting to see if unemployment has a different effect on males vs. females, youth vs. senior citizens, service industry vs. manufacturing etc. Unfortunately, the feasibility of such a study is questionable due to the lack of necessary international data.

Summary/Conclusions

The study initially set out to find a link between unemployment and suicide using a number of other socio-economic variables. As such, the results were more or less anticipated, with a possible exception: evidence was found to suggest that economic growth does not help to explain suicide rates, even during shorter periods of time. This contradicts a couple of research papers, including those of Andres (2005) and Neumayer (2004). This contradiction possibly stems from the difference in approaches used across the board.

What was not expected is the way income level affects the relationship between unemployment and suicide. Investigation suggests that the notion of suicide being positively dependent on unemployment is not exactly true. According to the analysis, unemployment has a positive significant effect on suicide rates in richer countries, while the opposite is true for poorer countries. It only goes to highlight that "what you've never had you never miss". Losing employment in the richer countries is much more distressing than in poorer countries.

The fact that suicide rates are sensitive to income levels contradicts the study of 15 European countries by Andres (2005). However, as was mentioned before, the effect of income on suicide rates would possibly be insignificant when the cross-section of the panel data consists of countries with relatively 'similar' income levels.

Overall, this paper should act as a guide for the researcher wishing to undertake a more comprehensive study of different factors affecting suicide rates (possibly when data are more abundant) and as a point of reference for the policy maker wishing to implement new health policies.

---

3 By trial and error, the threshold between high income and low income was found to be around US$ 7500.

References

Andres, A. R., 2005. Income inequality, unemployment, and suicide: A panel data analysis of 15 European countries. *Applied Economics*, 37, pp.440–450.

Chaloupka, F. J., Grossman, M., and Saffer, H., 2002. The Effects of Price on Alcohol Consumption and Alcohol-Related Problems. *Alcohol Research & Health*, 26(1), pp.22–34.

Chuang, H-L. and Huang, W-C., 1997. Economic and social correlates of regional suicide rates: A pooled cross-section and time-series analysis. *Journal of Socio-Economics*, 26, pp.276–290.

Department of Public Health on Behalf of the Chief Executive Officers of the Health Board, 2001. Suicide in Ireland. A National Study. Dublin: Department of Public Health.

Dorland's Illustrated Medical Dictionary, 32nd ed. 2011. Philadelphia, Pa.: Elsevier/Saunders, c2012.

Gerdtham, U-G. and Johannesson, M., 2003. A note on the effect of unemployment on mortality. *Journal of Health Economics*, 22, pp.500–520.

Neumayer, E., 2004. Recessions lower (some) mortality rates: Evidence from Germany. *Social Science and Medicine*, 58, pp.1035–1050.

United Nations, Department of Economic and Social Affairs: Population Division, Population Estimates and Projections Section Database, 2011. [Online] Available at: http://www.un.org/esa/population/ [Accessed 23 February 2011].

United Nations, Department of Economic and Social Affairs: Population Division, 2011. World Fertility and Marriage Database 2008 (incl. estimates). [CD-ROM] New York: UN.

World Bank Database, 2011. [Online] Available at: http://databank.worldbank.org/ddp/home.do [Accessed 23 February 2011].

Wooldridge, J. M., 2003. *Introductory econometrics: A modern approach.* 2nd

ed. USA: Thomson.

World Health Organization, Department of Mental Health and Substance Abuse, 2011. Global Status Report on Alcohol and Health 2010: Country profiles. [CD-ROM] Geneva: WHO.

World Health Organization, Mental Health Statistics, 2011. [Online] Available at: http://www.who.int/topics/mental_health/en/ [Accessed 23 February 2011].