Disentangling the associations between employment, income and antidepressant use: an application of the parametric G-formula

Maarten J. Bijlsma[a], Lasse Tarkiainen[b], Mikko Myrskylä[a,c], Pekka Martikainen[b]

[a] Max Planck Institute for Demographic Research, Rostock, Germany
[b] University of Helsinki, Department of Sociology, Helsinki, Finland
[c] London School of Economics & Political Science, London, United Kingdom

Introduction

Health has strong socio-economic disparities (Adler & Newman, 2002). Among socio-economic factors, employment and income are prominent, as they largely determine ones social and economic status, and hence affect health and mortality (Duncan et al., 2002). The effect of these factors has been widely studied, however modeling the causal effect of employment and income on health is troubled by endogeneity; employment and income may not just affect health, but also vice versa. Healthy individuals are more likely to become employed, and to perform better and thereby receive a higher income, compared to less healthy individuals (Wagenaar et al., 2012). Therefore, determining the effect of employment and income on health requires a careful disentangling of the time-varying reciprocal relationships between these three variables. Unfortunately, in the existing literature, these relationships are typically considered in a more static temporal way. This hinders estimation of the causal effects of these variables on one another, thereby potentially misinforming policy.

The parametric G-formula is a recently developed technique that allows for the disentangling of effects of time-varying variables (Keil et al., 2014). The technique is firmly grounded within the causal inference approach (Pearl, 2000), which can be seen as an improvement over structural equation modeling (Pearl, 2010). The technique can be used to control for measured time-varying confounders that are also affected by prior exposures (i.e. intermediate confounders), which standard regression approaches cannot accomplish. Importantly, the technique can be used to determine the effects of interventions (e.g. policy effects on employment or income) at the population level.

The aim of this study is to determine the effect of employment on antidepressant prescribing in Finland in the period 1996-2007, while taking into account time-dependent reciprocal relationships, and the mediating effects of income, education, and physical health.

Data and methods

Setting and data source

We extracted individuals from the Finnish EKSY data file who were aged 14 to 24 years old in Finland in the calendar period 1996-2001 and followed them from the moment they entered the labor market, until the year 2007. The Finnish EKSY data file is an 11% random sample of the population permanently residing in Finland at the end of any of the years in the period 1987-2007, and contains individual-level linked information on labor market records, census records, death records, social care records, medication records, sickness absence allowance records, and more.

Outcome variable, time-varying and time-invariant covariates

The outcome variable of interest is time from entering the labor market to first antidepressant prescribing. Time is measured in calendar years. The start of follow-up occurs when an individual enters the labor market. We identified 42,172 individuals who met these criteria. Follow-up ends if an individual receives at least one anti-depressant in a calendar year.

The time-varying covariates are education, income, employment status, household status, drug utilization (other than anti-depressants), age and calendar time. These variables are measured once a year. The majority of these variables are categorical. Time invariant covariates are parental income when individual lived at home, ethnicity and sex.

The G-formula

The G-formula works in three stages (Keil et al., 2014). In the first stage, we modeled conditional relations between variables allowing for a large number of possible relations, including lagged versions of variables as for such variables the time order can be more strongly ascertained. Model pruning, using a backward selection procedure ($p < 0.05$) resulted in some relations not being included in the final conditional models. Categorical variables were modeled with logistic regression and continuous variables with linear regression (including parametric curves).

In the second stage, we generated a 'natural course' scenario, and a number of counterfactual scenarios. The natural course scenario was generated as follows: We sampled individuals at the first year of follow-up. Then, covariate values in year 2 were predicted based on their covariate values in year 1 using the effect estimates from the first stage. The counterfactual scenarios are identical, except that simulated interventions take place to determine the employment effect, and mediation of the employment effect (Table 1).

In stage three, we contrasted the full intervention scenario and the mediation scenario with the natural course scenario. This is done by applying a Cox regression model with only an indicator variable for intervention (or mediation) scenario vs. natural course data as covariate (Keil et al. 2014).

Scenario	Variables generated as if unemployed = employed	To quantify
Natural course	No variables	-
Mediation (1)	AD prescribing	Direct effect of Employment
Mediation (2)	As (1) + Income	Employment effect via Income
Mediation (3)	As (2) + Drugs	Employment effect via Drugs
Mediation (4)	As (3) + Education	Employment effect via Education
Full intervention	All variables	Full employment effect

Table 1. The scenarios simulated through the G-formula. All mediation scenarios are contrasted with the natural course scenario.

Preliminary results

In the empirical data, at the first year of follow-up, 79% of individuals were employed and 21% were unemployed (start to follow-up required either employment or unemployment). In the first year, the majority of individuals were aged 20 years. These (and other) characteristics diversified during follow-up. Over the entire study-period, 75% of observed person-years were employed, 8% unemployed, 12% studying, and the remainder of ‘other’ status. The number of individuals in our closed cohort that received a first anti-depressant during the follow-up was 4320, or 12% of all individuals (see also Figure 1). These time-varying characteristics were matched by the natural-course scenario, indicating that the G-formula adequately modeled the empirical data (Table not shown due to space constraints).

G-formula

By intervening on employment so that all unemployed person-years were employed, the population-averaged hazard ratio between the natural course and the full intervention scenario was 0.942 (95% CI 0.906 to 0.985), indicating a 5.8% decrease in the hazard of antidepressant prescription. About 67% of this total effect is due to employment alone, 20% is due to income

5.6% is from drug utilization and 6.7% is due to education. The effect via household type accounted for less than 1%.

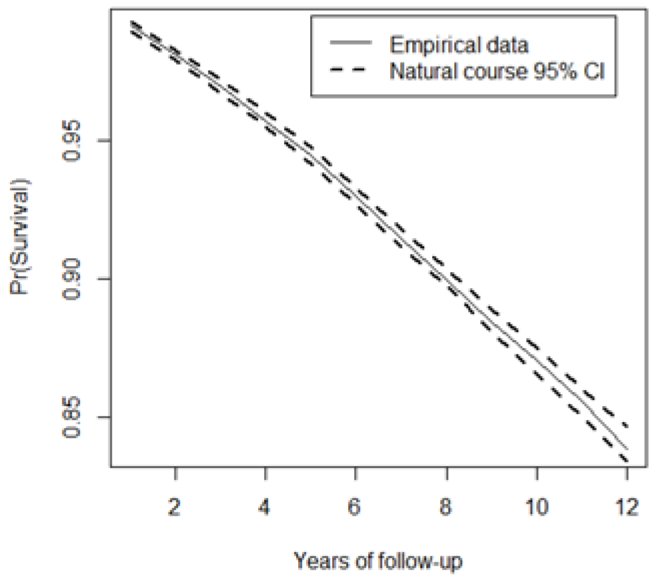

Figure 1. Survival function (Kaplan-Meier) from entering the labor market to first receiving a first antidepressant in the observed data together with 95% confidence intervals from the 'Natural Course' (no intervention) G-formula scenario.

Preliminary conclusion

The findings of this study indicate that being employed strongly lowers the risk of depression. These effect of unemployment is expected to be largely psychological, given that variables which are strongly associated with material living conditions accounted for about 1/3rd of the total effect of employment.

References

Adler NE, Newman K. Socioeconomic disparities in health: pathways and policies**.** Health Aff. 2002;21(2):60-76.

Duncan GJ, Daly MC, McDonough P, Williams DR. Optimal Indicators of Socioeconomic Status for Health Research. American Journal of Public Health 2002; 92 (7): 1151-57.

Keil AP, Edwards JK, Richardson DB, Naimi AI, Cole SR. The parametric g-formula for time-to-event data: intuition and a worked example. Epidemiology. 2014; 25(6):889-97.

Pearl J. Causality: Models, Reasoning, and Inference. Cambridge: Cambridge University Press; 2000.

Pearl J. The Foundations of Causal Inference. Sociological Methodology 2010; 40:75-149.

Wagenaar AF, Kompier MA, Houtman IL, van den Bossche SN, Taris TW. Employment contracts and health selection: unhealthy employees out and healthy employees in? J Occup Environ Med. 2012; 54 (10):1192-200.