Social Networks 40 (2015) 55–63

Contents lists available at ScienceDirect

Social Networks

journal homepage: www.elsevier.com/locate/socnet

The effects of specific occupations in position generator measures of social capital

Martin Hällsten[a,*], Christofer Edling[b], Jens Rydgren[a]

[a] Department of Sociology, Stockholm University, Sweden
[b] Department of Sociology, Lund University, Sweden

ARTICLE INFO

Keywords:
Position generator
Jack-knife
Social capital
Measurement error

ABSTRACT

The position generator is a widespread method for measuring latent social capital in which respondents are queried about contacts on a list of occupations predefined by the analyst. We separate out the unique contribution of each occupation to aggregated measures of social capital. It turns out that this contribution varies vastly: knowing a person in some occupations provides substance to measures of social capital, while knowing a person in a few occupations is irrelevant and contributes statistical noise and causes attenuation bias. We discuss the implication of our findings for the design of position generator measures generally.

1. Introduction

Social relations between individuals are important for understanding many processes in contemporary society. Such relations are often referred to as *social capital*, and they are an increasingly popular explanation for social phenomena such as political participation, civic engagement, population health and longevity, and labor market outcomes. Sociologists such as Bourdieu (1986), Lin (2001, 2002), and Burt (2005) view social capital as individual-level resources (e.g. monetary resources) or other valued assets (e.g. information) that individuals can access through their social networks and use to achieve positive (or cause negative) outcomes. For our purposes, we will follow Lin (2001, 2002), who see social capital as formed out of both social embeddedness and strategic behavior (i.e., investment behavior).

The position generator (Lin and Dumin, 1986) offers a coherent measure of individual-level social capital. The instrument is simple. Individuals are queried about contacts with persons in different occupations, using a pre-specified list chosen to represent positions in the labor market. The list of occupations varies across implementations of the position generator, although many follow Lin and Dumin's seminal list. Previous research has used position generator social capital to understand, among other things, civic engagement (Magee, 2008), marriage patterns (Lai, 2008), and psychological well-being (Moore et al., 2009). Most research, however, is focused on explaining inequality in labor market outcomes (Erickson, 2001; Flap and Boxman, 2001; Lin and Ao, 2008; Lin et al., 2001), which will also be our focus in this paper.

The aim of the position generator is to produce a *generic measure* of social capital, which can be used to explain subsequent outcomes. However, recent findings suggest that the position generator may not be as generic or content-free as desired (Van der Gaag et al., 2008, 2012; Verhaeghe et al., 2012). We add to this literature by analyzing whether the measure may be contingent on the analysts' choice of specific occupations that are included in the position generator. In order to evaluate each of the occupations' contribution to the measure of social capital, we set up a jack-knifing procedure in which criterion outcome variables are regressed on measures of social capital, where each of the occupations in the position generator list is in turn removed.

The paper is organized as follows: we briefly discuss the problems involved in measuring social capital and describe the position generator instrument. We then outline our analytical jack-knifing procedure, present our results, and close with a discussion of the practical and theoretical consequences for both the measurement and the concept of social capital itself.

2. Measures of social capital

Measuring individual-level social capital is difficult. The overall problem is that social capital often becomes visible to analysts only *after* it has been activated and employed through agency, yet there are many situations in which action is not necessary. Social capital is latent, it is mobilized only when needed, and observed only when

* Corresponding author. Tel.: +46 8 163556.
E-mail address: martin.hallsten@sociology.su.se (M. Hällsten).

http://dx.doi.org/10.1016/j.socnet.2014.06.002
0378-8733/

mobilized. This is different from most forms of resources such as human capital: an individual's education, experience, and tenure are often readily observable by fairly simple methods.

Much of the literature implicitly or explicitly gauges social capital as *ex post realizations*, for example by asking how a person got a particular job, where personal contacts is one obvious response. Measures of *mobilized* social capital are always conditioned on a situation where mobilization was necessary, and often on mobilization producing some positive outcome. Asking individuals how they got their job ignores individuals who fail to get a job (with or without mobilizing social capital). Ex post realization captures the *effects* of social capital, not the social capital resource itself, which is tautological (as critically discussed in Fernandez and Fernandez-Mateo, 2006; Lin and Dumin, 1986). Retrospective measures can therefore produce highly misleading results.

An ideal measure of social capital should tap the latent utility of social contacts *before* they are activated. The position generator is appealing for this reason. In a survey setting, respondents are asked if they know someone on a list of occupations, and then to further qualify whether this is a strong or weak contact (Lin and Dumin, 1986). Common response alternatives are that contacts are friends, acquaintances, or kin. This survey procedure allows the researcher to get a picture of the respondent's contact network and potential network resources. However, as discussed by Lin and Dumin, it is not feasible to ask for all occupations that exist, and the procedure is largely blind to the exact qualitative value of each specific contact (although tie strength is commonly queried). The measures generated out of the position generator thus *proxy* for direct network effects such as transfer of information, influence etc., and can be used to explain subsequent outcomes. Hence, the position generates gains in causal order but can also be said to lose in specificity and detail since the active mechanisms will remain latent.

For each individual, the position generator produces discrete information on access to each of the occupations queried for. This information can be further processed and refined by using characteristics on the specific occupation, where the gold standard is to measure the value of occupations in terms of their prestige or status. For each dimension, this information is then summarized *within* individuals. Lin et al. (2001) proposed computing the following dimensions of social capital:

- Extensity (total number of occupational contacts)
- Upper reachability (the highest accessed prestige)
- Prestige range (the range between highest and lowest access prestige)

Van der Gaag et al. (2008) also discussed some common complementary measures:

- Average prestige (average prestige per contact)
- Total prestige (total accessed prestige)

These measures are often highly correlated, and to get a unidimensional measure of social capital, Lin et al. (2001) have suggested using factor analysis to compute a composite measure that overcomes these problems.[1] Even though the underlying data structure is truly discrete, the first step of summing over all the occupations for each individual creates at least approximate continuous variables that are suitable for factor analysis. While this generic measure conceals differences in the underlying dimensions it circumvents the problems of multicollinearity, which can be severe in small samples.

3. Previous labor market studies employing the position generator

There is now consistent evidence that social network resources—using different instruments—are positively correlated with labor market outcomes such as job prestige, income, and wages (Lin, 1999).[2] In 1986, Lin and Dumin (1986) introduced the position generator methodology and showed that there was a great deal of inequality in the access to social capital. In a comparison of different measures of social capital, Van der Gaag et al. (2008) argued that the position generator tends to reflect resources useful in instrumental (as opposed to expressive) actions. A number of studies have shown that position generator social capital is positively associated with labor market outcomes (Erickson, 2001; Flap and Boxman, 2001; Lin and Ao, 2008; Lin et al., 2001).

Since Lin and Dumin's seminal study, a substantial number of studies have employed the position generator instrument. Bartelski (2010) listed 42 surveys that include the position generator. Many of them have small samples size (<1000) and only include 10–20 occupations, but there are exceptions. Many of these have examined access to social capital, labor market returns on social capital, and the consequences of social capital for other outcomes such as trust and political participation. This is not the place to summarize all of these studies, although the majority of them find non-trivial associations between position generator social capital and labor market outcomes.

4. Data

The position generator was included in the 2009 wave of the Swedish survey *Social Capital and Labor Market Integration*, which is a survey of school leavers. The survey sample consisted of 5695 individuals selected for telephone interview and carried out by Statistics Sweden between October and December 2009. A total of 2942 interviews were conducted, hence a response rate of 51.6%. The sample was based on three different cohorts of Swedes born in 1990: (a) all individuals with at least one parent born in Iran; (b) 50% of all individuals with at least one parent born in (former) Yugoslavia; and (c) a simple random sample of 2500 individuals with two Swedish-born parents. It thus gauges conditions for young individuals at age 19.

In the survey, a list of 40 occupations represents the contact space.[3] Table 1 shows the occupations and their associated social class codes (Swedish SEI, Statistics Sweden, 1982)—analogous to the Erikson-Goldthorpe-Portocarero EGP scheme (Erikson and Goldthorpe, 1992)—, the ISCO code, Treimans occupational prestige (Ganzeboom and Treiman, 1996), the occupational size according to the 1990 census (number of incumbents), and finally the access to the occupation in the survey and its associated variance. Because access is a discrete Bernoulli variable; the variance is only a function of the mean ($\sigma^2 = p(1-p)$). We compute variance of access since the performance of the social capital measure depends on

[1] Lin et al. (2001) and Lin and Ao (2008) used the first *rotated* principal component factor, whereas Van der Gaag et al. (2008) suggested using the first *unrotated* factor of a factor analysis. The reasons behind these diverse practices are beyond the scope of the paper. In our case the rotated solution is less associated with our criterion outcomes than the unrotated solution; hence the latter was chosen for further analysis.

[2] One should note that the causation of this association is contested, homophily being proposed as an alternative explanation (Mouw, 2003).

[3] The wording of the survey question is as follows: "I will read you a list of occupations and ask you if a close friend, acquaintance, family member, girlfriend/boyfriend, or relative has the occupation. [Name of occupation]? Does any close friend, acquaintance, family member, girlfriend/boyfriend, or relative have that profession?

Table 1
Occupations in the *Social Capital and Labor Market Integration* survey position generator.

Item nr	Occupation	Social class[a]	ISCO-88(com)	Prestige[b]	Size[c]	Access[f]	Variance in access[e]
1.	MEDICAL DOCTOR	56	2221	78	24,751	0.38	0.24
2.	COOK	22	5122	31	31,028	0.51	0.25
3.	CONSTRUCTION WORKER	21	7129	28	40,844	0.57	0.25
4.	ASSISTANT NURSE	12	5132	42	167,555	0.56	0.25
5.	ENGINEER	56	2142	70	2634	0.38	0.23
6.	HAIRDRESSER	12	5141	32	23,488	0.61	0.24
7.	MAILMAN	12	4150	33	28,928	0.28	0.20
8.	LAWYER	57	2421	73	2553	0.20	0.16
9.	DISABLED'S ASSISTANT	12	5133	17	110,226	0.53	0.25
10.	FACTORY WORKER	21	8283	30	22,442	0.46	0.25
11.	TELEMARKETER	33	5227	32	4887	0.54	0.25
12.	TEACHER	46	2330	57	53,253	0.52	0.25
13.	NURSE	46	3239	44	52,834	0.42	0.24
14.	TRUCK DRIVER	12	8323	33	70,343	0.44	0.25
15.	ESTATE AGENT	46	3413	49	8825	0.13	0.12
16.	PROFESSIONAL MUSICIAN	56	2453	45	3326	0.26	0.19
17.	POLICE OFFICER	46	3450	45	17,506	0.32	0.22
18.	CLEANER	12	9122	21	111,119	0.38	0.24
19.	DENTIST	56	2222	70	8607	0.28	0.20
20.	MECHANIC	21	7231	43	28,917	0.51	0.25
21.	CHILD CARE ASSISTANT	22	5131	23	122,836	0.33	0.22
22.	SELF-EMPLOYED WITH STAFF	79	1314	46	65,557	0.61	0.24
23.	CASHIER STAFF	12	4211	34	22,970	0.64	0.23
24.	SECURITY GUARD	12	5152	30	10,879	0.32	0.22
25.	REPORTER	46	2451	58	23,293	0.18	0.15
26.	PROFESSIONAL ACTOR	46	2455	57	1995	0.08	0.08
27.	RECEPTIONIST	33	4222	38	15,246	0.24	0.18
28.	FINANCIAL MANAGER	56	1231	60	10,617	0.15	0.13
29.	UNIVERSITY STUDENT	56	–	56[*]		0.84	0.13
30.	TAXI DRIVER	79	8321	31	14,188	0.32	0.22
31.	HEADMASTER	57	1227	60	7207	0.17	0.14
32.	COMPUTER TECHNICIAN	46	3121	53	6297	0.44	0.25
33.	RECREATION LEADER	46	3462	49	16,030	0.33	0.22
34.	BANK CLERK	46	3418	46	39,971	0.23	0.18
35.	WAREHOUSEMAN	11	4131	30	65,485	0.46	0.25
36.	COMPUTER PROGRAMMER	56	2131	51	27,631	0.32	0.22
37.	ACCOUNTANT	60	2411	62	16,703	0.14	0.12
38.	CARETAKER/JANITOR/ATTENDANT	12	7137	25	41,704	0.24	0.18
39.	RESEARCHER	56	2211[d]	69	1157[d]	0.16	0.14
40.	WAITER/WAITRESS	22	5123	21	16,452	0.63	0.23

Notes:
[a] Coded to Swedish SEI class scheme (Statistics Sweden, 1982) which is equivalent to EGP class scheme (Erikson and Goldthorpe, 1992).
[b] Treimans SIOPS (Ganzeboom and Treiman, 1996).
[c] Number of incumbents according to 1990 census.
[d] These occupations can also be found in alternative ISCO codes.
[e] Variance = $p(1-p)$.
[f] Access = proportion of sample with the occupational contact.
[*] Assumed for a successful student.

the variance of its components. All else being equal, components with large variation will contribute more than components with lesser variation. A logical consequence of this is that occupations with access in proportion at around $P = 0.5$ will contribute the most (again, all else being equal, e.g., occupational prestige and resources tied to the occupation), since this is the point where variance is maximized.

The 40 occupations were chosen on the basis of earlier implementations (Lin and Dumin, 1986; Lin et al., 2001; Van der Gaag et al., 2008). They represent the range of class and occupational prestige in Swedish society fairly well, as judged by comparisons with census statistics on occupational prestige distributions and monthly wages (see Figures A1–A4 in the Appendix). However, compared to reference occupations in the population, the 40 occupations to a somewhat larger degree represent the extremes. Occupations with low prestige and low average wages are oversampled in order to increase variance; knowing individuals in less common occupations will provide information with higher discriminatory value. This also holds for occupations with high prestige, and to some extent for occupations with high wages as well.

It should be noted that our population is special in that it is focused on young individuals on the verge of labor market entry. This has consequences for our study, and is both an advantage and disadvantage. Even though we can eliminate heterogeneity due to age variation, the list of 40 occupations may at first glance be more accurate for the general population than for 19-year-olds. However, since individuals are in the process of labor market entry, few of their contacts are structured by their own independent labor market career, but are rather structured by their parents' socio-economic position.

The social capital measure we use in this paper is the first unrotated principal factor, which uses all of the dimensions discussed above (extensity, upper reachability, range of accessed prestige, average prestige, and total prestige). This composite specification was chosen in part because it is common in the literature, and in part because it explains most of the criterion variables (described below).[4] Table 2 shows the eigenvalues and factor loadings of our composite measure.

[4] The only other Swedish application of the position generator that we are aware of uses extensity, upper reachability, prestige range, and average prestige to generate a composite generic measure of social capital (Behtoui, 2007). The present paper is a part of a Swedish project that aims to analyze the role of social relations for young adults' life chances.

Table 2
Factor analysis for baseline composite measure of social capital.

	Five components			Three components	
Eigenvalues					
Factor1	3.395			2.186	
Factor2	0.985			0.079	
Factor3	0.376			−0.084	
Factor4	−0.006				
Factor5	−0.037				
Factor loadings	(1)	(2)	(3)	(1)	(2)
Extensity (number of positions accessed)	0.834	−0.536	0.103	0.629	0.227
Range of prestige accessed	0.892	0.127	−0.358	0.967	0.012
Upper reachability (highest prestige)	0.888	0.365	−0.194	0.925	−0.166
Average prestige	0.558	0.634	0.396		
Total prestige	0.897	−0.384	0.206		
N	2891			2891	

Note: Principal factor solution.

Table 3
Descriptive characteristics.

Short label	Description	Mean/(SD)	Min/Max	*N*
Alter	Alter's characteristics (canonical factor)	2.43 (1.000)	−0.714 6.057	2664
GPA	GPA (upper-secondary)	12.894 (5.081)	0 20	2227
UnE	Unemployment (0/1)[a]	0.341 (0.474)	0 1	1590
SB	Social background (principal component factor)	0 (0.909)	−2.81 4.755	2581
Ln CPD	Ln contacts per Day	3.333 (1.003)	0 6.215	2873

Note:
[a] Delimited to those in risk = not in education.

5. The jack-knife procedure

The aim of this paper is to assess the unique contributions of each occupation included in the position generator inventory of occupations. In order to do this, we set up a jack-knifing procedure. We evaluate the association between the composite social capital measure as based on different occupations out of the list of 40, and criterion variables, that is, outcome and background factors that we *presume* should be associated with social capital. Here we consider characteristics of up to five alters' from a name generator (Alter),[5] grade-point average from upper-secondary school (GPA) from which they just graduated, own unemployment if not studying (UnE), a social background index (SB),[6] and (the natural log of) the number of persons ego has contact with each day (Ln CPD).[7] Table 3 shows descriptive statistics for these variables.

The jack-knifing methodology is straightforward and can be described as follows. First, we use all of the underlying dimensions variables (extensity, upper reachability, range of accessed prestige, average prestige, and total prestige) in a factor analysis to compute a baseline measure of generic social capital, SC, using all 40 occupations. We then run a regression of the criterion variables, Y, on social capital, using all occupations.

$$Y = a + b \times \mathrm{SC}_{\mathrm{BASELINE}} + e \tag{1}$$

This we regard as our baseline model. We then create jack-knifed versions of the underlying dimensions by removing each of the 40 occupations (j) one-by-one before summarizing the components within individuals. Information from all five dimensions is used to create 40 measures of generic social capital, using factor analysis. We then run 40 univariate regressions of the criterion variables on the jack-knifed social capital measures

$$\text{for } j = 1 \text{ through } 40: \quad Y = a + b_{-j} \times \mathrm{SC}_{-j} + e \tag{2}$$

Our target parameters for the jack-knife procedure are t and R^2, and so we record the difference in t-values and R^2 between the 40 jack-knife regressions and the baseline regression. The t-value is the regression coefficient scaled by its standard error and captures both the occupation's effect on the regression coefficient and its statistical significance. We also include effects on R^2, which estimates the effect on the overall explained variance of social capital. The t and R^2 values are related and should be seen as different scales rather than dimensions.[8] Eqs. (3a) and (3b) describe how we achieve our jack-knifed measures of the influence of one occupation on t and R^2:

$$\Delta t_j = t(b)_{\mathrm{BASELINE}} - t(b_{-j}) \tag{3a}$$

$$\Delta R_j^2 = R_{\mathrm{BASELINE}}^2 - R_{-j}^2 \tag{3b}$$

Positive Δt means that t-values *decrease* when the occupation is *removed*, which implies that the occupation contributes information that is more important than the average occupation. *Negative* Δt means that t-values *increase* when the occupation is removed, meaning that the occupation contributes information that is less important than the average occupation.

However, removing occupations one by one does not fully disclose how large the impact of the choice of occupations can potentially be on the results. Choosing different *sets* of occupations (within the list of 40 occupations) may have a larger impact. In a second step, we repeat the jack-knife analyses by removing instead sets of 10 randomly chosen occupations. We use 200,000 replications in order to capture all possible combinations. With this procedure, we can identify those combinations of occupations that contribute most to the measure of social capital.

[5] We measure alters' characteristics using a canonical correlation model. We extract the factor that maximizes the correlation between social capital and an array of alter characteristics and use this as our measures of alters. The alter characteristics we use are share of unemployment among friends, proportion of university entrants, smoking habits, exercising behavior, food habits, over-weight, ego's trust in alters and ego's judgment of alter's risk behavior. The extracted canonical factor correlates .2745 with the measure of social capital.

[6] This is the first unrotated factor of the following variables (measured on both mothers and fathers): years of education, employment (0/1 coded), occupational prestige, and disposable income. This information comes from administrative registers.

[7] We have also experimented with the number of nominated alter non-kin friends, the number of unemployed alters, and the average homogeneity in terms of alters' age, smoking, risk taking behaviors, etc., but none of them proved to have any strong relation to social capital and were therefore less useful as criteria.

[8] Bring (1994) shows that t-values and R^2 are functions of the same latent factor.

Table 4
Jack-knife estimates of differences in *t*-values (Δt) and *R*-squared (ΔR^2) for social capital with single occupations removed by outcomes/background factors.

	Δt					ΔR^2				
	Alter	GPA	UnE	SB	Ln CPD	Alter	GPA	UnE	SB	Ln CPD
1 Medical Doctor	0.48	−0.11	0.23	0.74	0.25	0.45	−0.06	0.22	0.41	0.18
2 Cook	−0.04	−0.05	0.06	−0.06	0.05	−0.04	−0.03	0.06	−0.03	0.04
3 Construction Worker	0.05	0.14	−0.07	0.12	0.12	0.04	0.08	−0.07	0.07	0.09
4 Assistant Nurse	−0.03	0.04	0.03	−0.14	−0.02	−0.03	0.02	0.03	−0.08	−0.01
5 Engineer	0.31	0.90	0.31	1.08	0.05	0.29	0.48	0.29	0.59	0.03
6 Hairdresser	−0.07	−0.06	0.02	−0.17	0.01	−0.06	−0.03	0.01	−0.10	0.01
7 Mailman	0.02	0.03	0.02	0.09	0.03	0.02	0.02	0.02	0.05	0.02
8 Lawyer	−0.09	−0.21	0.08	−0.39	0.01	−0.09	−0.12	0.08	−0.23	0.01
9 Disabled's Assistant	−0.02	−0.06	0.07	−0.24	0.13	−0.02	−0.03	0.07	−0.14	0.09
10 Factory Worker	−0.07	−0.03	−0.06	−0.06	0.00	−0.07	−0.02	−0.06	−0.04	0.00
11 Telemarketer	0.03	0.02	0.04	−0.01	0.04	0.03	0.01	0.04	−0.01	0.03
12 Teacher	0.30	0.28	−0.20	0.10	0.00	0.28	0.16	−0.19	0.06	0.00
13 Nurse	0.05	0.07	0.09	0.15	−0.02	0.05	0.04	0.09	0.09	−0.01
14 Truck Driver	−0.11	−0.16	−0.03	−0.15	0.04	−0.11	−0.09	−0.02	−0.08	0.03
15 Estate Agent	0.05	0.03	0.04	0.06	−0.03	0.05	0.02	0.04	0.03	−0.02
16 Professional Musician	−0.03	0.01	−0.05	0.04	0.04	−0.03	0.01	−0.05	0.02	0.03
17 Police Officer	−0.03	−0.03	−0.06	0.05	0.09	−0.03	−0.02	−0.06	0.03	0.06
18 Cleaner	−0.02	0.03	−0.06	−0.02	−0.12	−0.02	0.02	−0.06	−0.01	−0.09
19 Dentist	−0.05	0.29	−0.21	0.13	0.04	−0.05	0.16	−0.20	0.07	0.03
20 Mechanic	−0.21	−0.27	−0.10	−0.15	0.08	−0.21	−0.16	−0.09	−0.09	0.06
21 Child Care Assistant	0.01	0.05	0.00	0.04	0.00	0.01	0.03	0.00	0.02	0.00
22 Self−Employed With Staff	−0.01	−0.02	0.07	−0.14	0.15	−0.01	−0.01	0.06	−0.08	0.11
23 Cashier Staff	0.04	0.10	0.09	0.03	0.01	0.04	0.05	0.09	0.02	0.01
24 Security Guard	−0.02	−0.11	0.00	−0.10	0.05	−0.02	−0.06	0.00	−0.06	0.04
25 Reporter	0.06	0.13	0.11	0.17	0.00	0.06	0.08	0.10	0.09	0.00
26 Professional Actor	−0.07	−0.04	−0.07	0.00	−0.07	−0.07	−0.02	−0.07	0.00	−0.05
27 Receptionist	0.07	0.07	0.06	0.09	0.04	0.07	0.04	0.06	0.05	0.03
28 Financial Manager	0.26	0.12	0.29	0.26	0.10	0.25	0.07	0.27	0.15	0.07
29 University Student	0.91	0.32	0.38	0.15	0.12	0.84	0.17	0.35	0.09	0.08
30 Taxi Driver	−0.02	−0.08	−0.03	−0.15	0.00	−0.02	−0.04	−0.03	−0.09	0.00
31 Headmaster	0.00	0.07	0.07	0.21	−0.05	0.00	0.04	0.06	0.12	−0.04
32 Computer Technician	−0.11	−0.19	−0.27	−0.22	0.14	−0.11	−0.11	−0.26	−0.13	0.10
33 Recreation Leader	0.05	−0.03	−0.04	−0.09	0.12	0.05	−0.01	−0.03	−0.05	0.09
34 Bank Clerk	0.05	0.09	0.07	0.18	0.06	0.05	0.05	0.07	0.11	0.04
35 Warehouseman	0.04	−0.11	−0.03	0.05	0.10	0.04	−0.06	−0.03	0.03	0.07
36 Computer Programmer	−0.06	−0.05	−0.14	−0.02	0.07	−0.06	−0.03	−0.13	−0.01	0.05
37 Accountant	0.12	0.20	0.19	0.01	−0.23	0.11	0.11	0.18	0.00	−0.17
38 Caretaker/Janitor/Attendant	−0.03	−0.01	0.00	−0.01	−0.02	−0.02	−0.01	0.00	0.00	−0.01
39 Researcher	0.42	−0.01	−0.02	0.76	−0.08	0.39	0.00	−0.02	0.42	−0.06
40 Waiter/Waitress	0.07	0.05	0.04	0.18	−0.04	0.07	0.03	0.04	0.10	−0.03
Maximum jack-knife	0.91	0.90	0.38	1.08	0.25	0.84	0.48	0.35	0.59	0.18
% of Baseline	6.2	14.0	4.7	14.0	2.2	11.1	25.8	8.7	25.9	4.2
Minimum jack-knife	−0.21	−0.27	−0.27	−0.39	−0.23	−0.21	−0.16	−0.26	−0.23	−0.17
% of Baseline	−1.4	−4.2	−3.3	−5.1	−2.0	−2.8	−8.6	−6.5	−10.1	−3.9
	t					***R*²**				
Baseline model (all 40 occupations)	14.73	6.45	8.15	7.7	11.36	7.54	1.86	4.02	2.28	4.31

Note: The estimated models is $Y = a + b \times \text{SC}_{-j} + e$, where SC_{-j} refers to social capital with each the 1–40 occupations removed, and *Y* is one of the following outcomes: Alter = canonical factor of alter's characteristics (weighted by correlation with social capital, see text), GPA = grade-point average from elementary school, UnE = unemployment (0/1), SB = principal component factor of social background indicators (see text), CPD = the number of persons ego has contact with every day.

6. Results

The results for each occupation are displayed in the upper panel of Table 4, with baseline reference values in the lower panel. Roughly half of the occupations contribute positive Δt, whereas the other half contribute negative Δt—this is an expected pattern given that social capital is constructed as an average across many occupations. (Remember that Positive Δt means that the occupation contributes positively and that estimated associations would be weaker if the occupation was dropped). In many variables, the differences are non-negligible: the *t*-value can change up to 14% if one single occupation is removed (the middle panel of Table 4 compares Δt to baseline *t* in). Likewise, ΔR^2 shows a similar pattern but in another metric. Here, one single occupation can amount to as much as 25% of the explained variance of social capital (see middle panel of Table 4). In relative terms, these differences are large.

We make a few additional observations. For each occupation, the contributions (Δt and ΔR^2) to the social capital measure vary across outcome criteria. For example, knowing an engineer (occupation #5) contributes .90 *t*-units to explaining GPA, around 14% of the baseline *t*-value of 6.45. Knowing an engineer is also a key component for the social capital measures association with social background (more than 1 *t*-units), the quality of alters, unemployment risk and contacts per day. Other occupations contribute differently. Knowing a mechanic (occupation #20) is negatively associated with the social capital measures capacity to explain GPA, social background, and unemployment risk, but weakly positively correlated with the capacity to explain contacts per day.

In Table 5, we quantify how the social capital and criterion variable correlation is similarly affected by an exclusion of an occupation from the position generator. We thus analyze correlations of Δt for each of the outcomes across all occupations (that is, we analyze the correlations between columns in Table 4 for rows #1–#40).

Table 5
Correlations in Δt across occupations.

	Alter	GPA	UnE	SB	Ln CPD
Alter	1				
GPA	0.49	1			
UnE	0.62	0.42	1		
SB	0.62	0.64	0.44	1	
Ln CPD	0.15	−0.16	0.05	0.04	1

Note: the correlations is calculated across the 40 occupations as units of observations (see Table 4).

Since the correlations in ΔR^2 are close to identical, we omit these. The correlations are in most cases positive and substantial except for contacts per day (Ln CPD), which appears to be somewhat of an independent dimension. However, the strongest correlations are not higher than .65, which leaves ample room for occupation-specific effects and patterns. With alternative criterion measures, capturing other dimensions of life than the socio-economic dimension in focus here, it is likely that the correlations would be different than those displayed here.

In the next step, we run similar regression analyses where we remove blocks of occupations. We depart from the list of 40 occupations, and then randomly remove 10 occupations (for logical reasons, without replacement). We repeat this 200,000 times. Fig. 1 shows the distribution of Δt for the five criterion variables. There clearly is variation in the importance of the 30 occupations chosen for inclusion. The mean of the distribution of Δt is above 0, often up to .5 t-units. This means that many of the 40 occupations do contribute substantively to the measure, and that removing occupations randomly would weaken the measure. We can make two further observations. First, the distribution has a substantial yet smaller part below zero. This indicates that the optimal measure need not include all 40 occupations in the list. The minimum of the distribution of Δt is up to 2 units below zero, indicating that some combinations of occupations contribute mainly statistical noise to the measure. Fig. 2 repeats the information but in the ΔR^2 measures. If we compare with the baseline R^2s as displayed in the bottom line of Table 4, it is evident that the choice of occupations matters a great deal. The extreme negative point in the distributions amounts to up to 50% of the explained variance (see also Table 6).

We dig further into this issue by recording the combinations of randomly removed occupations that yield the most negative and positive Δt and ΔR^2. Table 6 shows these occupational blocks and their associated parameters. For each criterion, the upper panel shows the combination that makes the most negative contribution. For example, for the Alter criterion, the t-value increases from 15.7 to 16.7 when the 10 listed occupations are removed. This corresponds to almost 1 percentage point in R^2 (close to 10% in relative terms). Hence, these occupations do not contribute to the measure of social capital. Even if the cost of including these occupations is small, nothing is gained by doing so. We discuss the reasons for why this is the case below. Many occupations recur on the list; for example, mechanic (occupation #20) and computer technician (#32) are included in the lists for alter characteristics, GPA, unemployment risk, and social background (but not for contacts per day, which was observed to be a disparate dimension in Table 5 above), but it is equally striking that the overlap has clear limitations.

Our statistical approach thus suggests that social capital is best defined without some occupations. It is also clear that some occupations are more pivotal than others for statistical power. The lower panel of Table 6 shows the most positive contributing blocks. For the Alter characteristics criterion, half or even more than half of the t-value and R^2 originates in only 10 occupations. If these were removed, we would severely underestimate the association. This is quite remarkable as it suggests that a social capital measure may be very sensitive to the inclusion of the right occupations in a position generator. In the field, most implementations of the position generator tend to draw on the original implementation by Lin and Dumin (1986), but there may of course be other occupations that have never been surveyed that may prove to have pivotal importance.

Examples of occupations that are more pivotal than others are medical doctors (occupation #1), engineer (#5), and financial manager (#28). Since we surveyed school-leavers, we included university students in the list (assuming that this will lead to high occupational prestige), and it turns out that knowing a university student (#29) is very important for social capital in many of the criterion variables. Hence, the importance of individual occupations for social capital has a skew distribution. We can also see patterns among those occupations that decrease the strength of the social capital measure. Mechanic (#20) and computer technician (#32) are such examples. It should be noted that there is less of a clear pattern among those occupations that contribute most negatively. In summary, the random selection simulation reinforces the conclusion that not all occupations are equally important.

We have made more systematic attempts to understand what drives the importance of an occupation by analyzing the contributions in Δt for each occupation and its correlation with three characteristics of occupations: (a) prestige of the occupation, (b) ln size of the occupation, and (c) variance of the average respondent access to the occupation (which is a function of the mean proportion). The results are displayed in scatterplots in supplementary Figures A5–A9. We cannot find any obvious pattern among our 40 occupations in these three dimensions. As the level of observation in this case is the occupations themselves, the lower number is a limit for meaningful analysis, but we can reject the hypothesis that some simple status or access dimension determines the statistical utility of an occupation in position generator measures of social capital.

Table 6
Occupations contributions to Δt and ΔR^2 assessed via random formation of blocks.

Outcome	Block	Baseline t	t	Δt	% of Baseline	Baseline R^2	R^2	ΔR^2	% of Baseline
Negative contributing occupation blocks:									
Alter	Removed: 04 07 14 15 16 20 22 32 35 36	14.729	15.837	−1.108	−7.520	0.075	0.086	−0.011	−14.379
GPA	Removed: 01 10 11 17 20 24 32 35 36 38	6.455	8.282	−1.827	−28.309	0.019	0.030	−0.012	−62.747
UnE	Removed: 12 16 18 19 20 21 27 32 33 36	8.152	9.374	−1.222	−14.984	0.040	0.052	−0.012	−30.603
SB	Removed: 02 08 09 10 14 20 23 30 32 36	7.699	9.494	−1.795	−23.309	0.023	0.034	−0.012	−50.497
Ln CPD	Removed: 08 12 13 15 18 30 31 35 37 38	11.362	12.153	−0.791	−6.965	0.043	0.049	−0.006	−13.709
Positive contributing occupation blocks:									
Alter	Removed: 01 05 08 09 15 22 24 28 29 39	14.729	9.705	5.024	34.111	0.075	0.034	0.041	54.570
GPA	Removed: 05 12 13 18 25 28 29 31 32 37	6.455	2.565	3.890	60.262	0.019	0.003	0.016	83.921
UnE	Removed: 01 05 08 09 15 22 24 28 29 39	8.152	5.397	2.755	33.801	0.040	0.018	0.022	55.109
SB	Removed: 01 05 12 13 16 19 25 30 38 39	7.699	1.684	6.015	78.126	0.023	0.001	0.022	95.108
Ln CPD	Removed: 01 05 09 10 17 29 32 34 36 39	11.362	9.566	1.796	15.806	0.043	0.031	0.012	28.068

Note: Based on 200,000 draws of 10 occupations randomly removed from original list of 40 occupations (sampled without replacement). See Table 1 for overview of occupations.

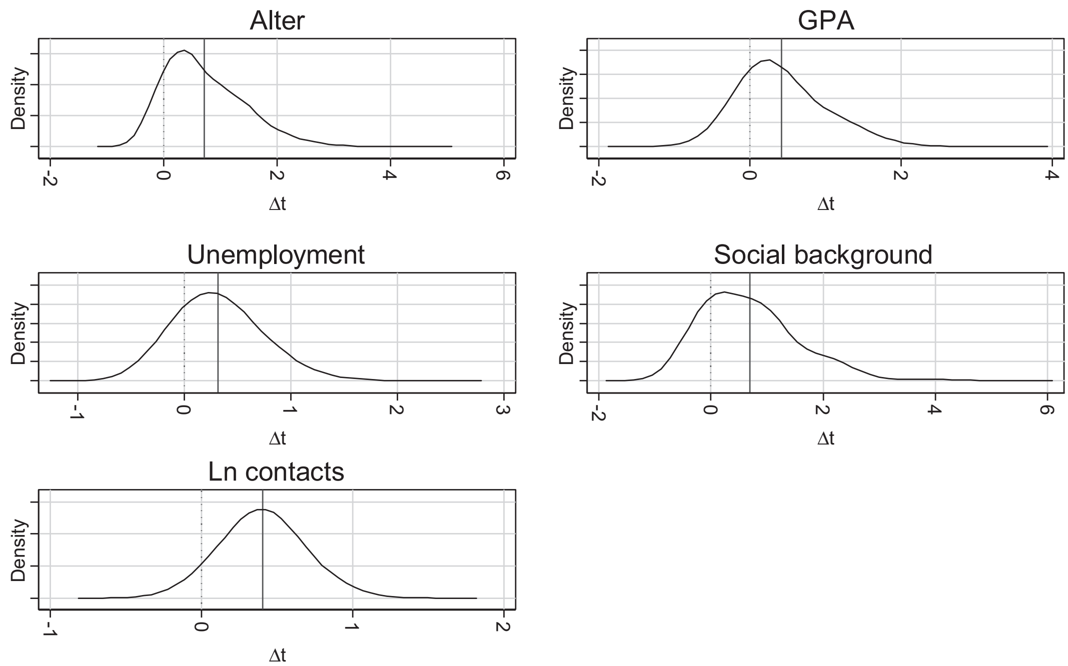

Fig. 1. Distribution of Δt removing random blocks of variables. *Note*: dotted line represents $\Delta t = 0$ and the solid line represents the mean of the distribution.

Finally, we have also done the analyses reported in Tables 4–6 for each of the component measures of social capital, instead of the composite index. Space does not allow discussion of all the details, which are available from the authors on request. Our results are *not* independent of which social capital measure we use. As judged by correlations in Δt (or ΔR^2), in most cases a common pattern dominates across occupations between different definitions of social capital for a specific criterion (similar to Table 5). This means that the occupation contributes similarly to extensity, upper reachability, prestige range, average prestige, and total prestige. However, there are some examples where a common pattern is lacking and the importance of specific occupations is unique to the component. This bolsters our conclusion that position generator social capital is contingent on the context and the research question.

7. Discussion

We have analyzed the validity of measures of social capital based on the *position generator* (Lin and Dumin, 1986), in which individuals are queried about contacts in a pre-specified list of occupations. Our analysis reveals that the composition of occupations in this list and/or in the measure of social capital is crucial to the effectiveness of the social capital measure. We show that some are pivotal. Omitting them would severely limit the statistical power of the social

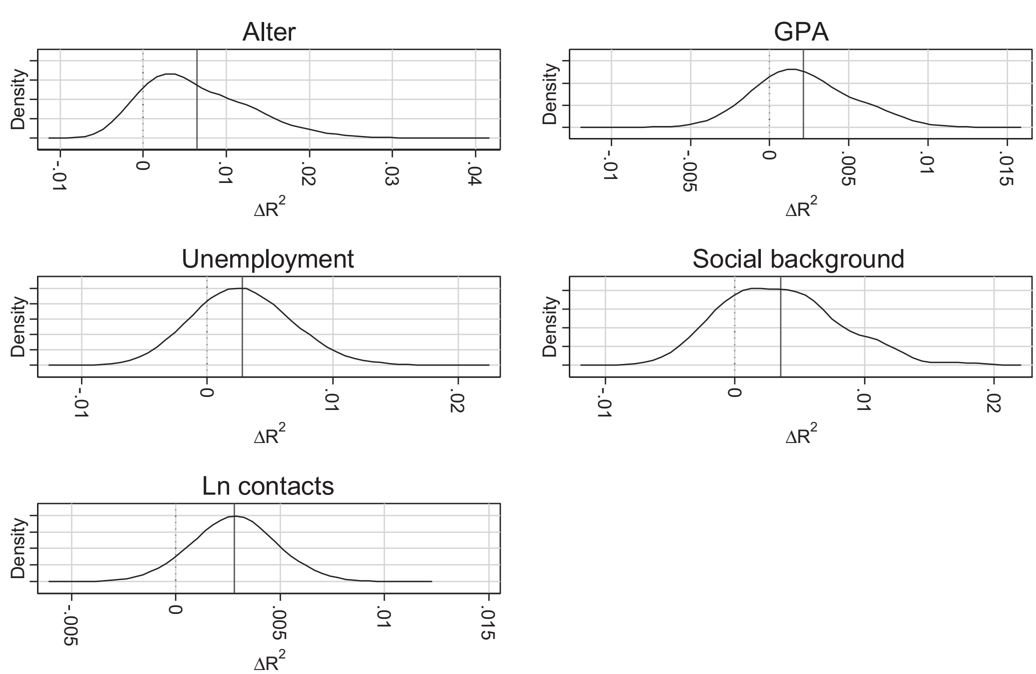

Fig. 2. Distribution of ΔR^2 removing random blocks of variables. *Note*: dotted line represents $\Delta R^2 = 0$ and the solid line represents the mean of the distribution.

capital measure; however, there are also some occupations that only contribute measurement error. Removing such occupations would positively impact the effectiveness of the social capital measure in predicting outcome variables. To some extent the logic is captured in the phrase *less is more*—by removing a few occupations, we can get a better measure. This runs against many recommendations in the literature. A big crux is that the positive or negative contribution of a measured occupational contact is very specific to the outcome process studied, including what population group we have at hand (youngsters, natives, immigrants) and social context and time (a post-industrialist developed welfare state in the early 20th century). Hence, measures should be adapted to the process under study rather than to be generic, which has the backdrop that comparability may suffer (or comparison may be tedious).

How can a measured occupational contact be negative for a measure of social capital? Obviously, knowing someone in those occupations is not consequential for the social process under scrutiny. The actual reasons for this can be manifold and specific to our population of school leavers, including incumbents' resourcefulness, control, generosity with help, and the like. A sample of 19-year-olds, as in our case, provides a specific context in which some relations, such as knowing a lawyer (see Table 2), are relatively unimportant. One can advance many theoretical arguments for why this importance should vary across the life-cycle, and knowing a lawyer may still be important at a later stage in life (e.g., for non-labor-market outcomes). Hence, apart from knowing which occupations proxy best for social capital, age-dependent variation in the importance of specific occupations may be a further confounding factor.

But how can this lack of usefulness be negative as we observed? We believe this can be explained by measurement theory. Measurement of independent variables with stochastic or classical errors is known to *attenuate* estimated effects. A very good illustration, which is somewhat analogous to our case, is that of intergenerational income mobility, where year-to-year variations in parents' earnings tend to depress the association between the earnings of fathers and sons, which can give highly misleading downward biased results (Solon, 1989). Hence, instead of measuring X, we measure $X + e$, and the regression coefficient of X becomes scaled down by a function of the variance of e (formally: b = true $b \times \sigma_X/[\sigma_X + \sigma_e]$). Some occupations add variation to our measure of social capital without contributing explanatory power. By including non-useful contacts in our measure, we actually get a negative effect on the effectiveness and power of our measure.

However, to the same extent that we have occupations that contribute negatively by inducing measurement error, when compiling the list of occupations that we query respondents about, we may have omitted occupations that may contribute strongly to social capital effects, that is, omitted variables bias.

Our conclusion is that the statistical power of composite social capital generated by the position generator is dependent on *specific* included occupations. Compared with other measures of social capital, the position generator was intended to be comparatively "content free," that is, independent of local context. However, the literature recognizes that this may not always be the case (Lin et al., 2001; Van der Gaag et al., 2008).[9] Our results indicate that the position generator may actually be dependent on context in a more fundamental way. Thus, in an ideal world, we should survey the utility of each and every occupation for social capital. However, this is a daunting and extremely expensive task. We cannot sample the whole potential contact space, yet the choice of sampling frame can be decisive for analytical precision. An appealing project would be to pool data from position generators included in studies in different nations and employ the jack-knifing technique to assess whether occupational contributions are stable across time, place, age, and outcome process. Without strong knowledge ex ante, it is difficult to assess whether an occupation will contribute noise or substance to any measure of social capital.

At this point, we can only provide some informed speculation about what is optimal in selecting occupations. First, when discussing measures of social capital, a central property that is important to keep in mind is the capacity to discriminate among individuals with high and low levels of resources embedded in social network relations. One may outline several mechanisms that make a certain occupation an important grid in this contact space. Uniqueness can be important because it may be associated with exclusive resources. For example, knowing the manager of a large company is more important if no one else knows this person, and we can expect information and influence to be more exclusive. We believe that this is related to occupational size: knowing someone in a highly common occupation is likely to provide resources, but unlikely to provide exclusive resources.

However, uniqueness may also be counterproductive for the efficiency of a measure of social capital (i.e., for achieving as much power within strict time and budget constraints) because variation will be small. For Bernoulli (dummy) variables, such as knowing someone in an occupation, variance is maximized when the proportion p is closest to .5, and considerably smaller when p is close to 0 or 1 (since $\sigma^2 = p(1 - p)$). Hence—while ignoring measurement error—occupations with more even distributions of access may be more viable as they contribute more variance, that is more discriminatory capacity, to measures of social capital. Concentrating the position generator only on occupations that are evenly distributed may, however, miss those opportunities plagued by the severe social closure that we pointed to above, so there is an obvious trade-off involved.

Second, apart from uniqueness and commonness, different occupations may be associated with different mechanisms of social capital. Lin (2001) distinguished between influence, information, and social credentials and reinforcement as different mechanisms of social capital. Accessing and indirectly exerting social influence often requires contact with a position of relative power. In our case, this may for example apply to the following occupations of higher than average importance: doctor (occupation #1), engineer (#5), financial manager (#28), and university student (# 29), which rank very high for most of our criterion variables (see Table 4). Information, on the other hand, may be more contingent on occupations that bridge structural holes Burt (1992), which are probably different from occupations of power. Examples from our study include bank clerk (# 34), receptionist (#27), and waiter (#40)—all of which contribute positively or above average in Table 4. Furthermore, occupational contacts useful for providing social credentials such as job referrals are likely to be of another kind, in most cases similar to the ego's human capital endowments or occupational specialization. Here it is difficult to make a general case for certain specific occupations.

The key conclusion of this paper is that measurement of social capital is more difficult than has been previously assumed. Different implementations of the position generator carry a contingent measurement error that depends on the context and process under study. The estimated effects of social capital may therefore be biased upward or downwards, but the direction of the bias is as specific as the measurement error. In most cases, our finding that some occupations contribute only noise lead us to suspect that social capital effects generally have been underestimated. Of course, earlier literature has been aware of some limitations and acknowledged that the position generator is only a proxy. However, in order to understand occupational contact networks, we would need more

[9] This may apply less to the measure of *extensity* of occupational contacts, which is presumably the most context-free of all dimensions in the position generator.

data on occupations. For any data collection, the number of sampled occupations should be as large as possible to make certain that no dimension is left out. The empirical measure of social capital, however, should use only those occupations that contribute more substance than measurement error to the measure for the given context, group and process under study, which in the present case has been found to be quite a low number. Without this optimization, attenuation bias is likely in the measure.

Acknowledgments

We are grateful to the Swedish Council for Working Life and Social Research (grant no. 2007-0806) and the Stockholm University Linnaeus Center for Integration Studies for supporting the data collection, and to the European Research Council for a Starting Grant awarded to Rydgren (grant no. 263422). The authors thank Pär Bendz for assistance with data, Love Bohman for comments on an earlier draft, and two anonymous referees.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.socnet.2014.06.002.

References

Bartelski, A.S., 2010. Social capital position generator studies. http://gaag.home.xs4all.nl/work/PG_studies.pdf
Bourdieu, P., 1986. The forms of capital. In: Richardson, J.G., Westport, C.T. (Eds.), Handbook of Theory and Research for the Sociology of Education. Greenwood Press.
Bring, J., 1994. How to standardize regression coefficients. The American Statistician 48, 209–213.
Burt, R.S., 1992. Structural Holes. Harvard University Press, Cambridge, MA.
Burt, R.S., 2005. Brokerage and Closure An Introduction to Social Capital. Oxford University Press, Oxford.
Erickson, B.H., 2001. Good networks and good jobs: the value of social capital to employers and employees. In: Lin, N., Cook, K., Burt, R.S. (Eds.), Social Capital: Theory and Research. Aldine de Gruyter, New York, NY, pp. 127–158.
Erikson, Robert, Goldthorpe, J.H., 1992. The Constant Flux: A Study of Class Mobility in Industrial Societies. Clarendon Press, Oxford.
Fernandez, R.M., Fernandez-Mateo, I., 2006. Networks, race, and hiring. Am. Sociol. Rev. 71, 42–71.
Flap, H., Boxman, E., 2001. Getting started: the influence of social capital on the start of the occupational career. In: Lin, N., Cook, K., Burt, R.S. (Eds.), Social Capital: Theory and Research. Aldine de Gruyter, New York, NY, pp. 159–181.
Ganzeboom, H.B.G., Treiman, D.J., 1996. International Stratification and Mobility File: Conversion Tools. Department of Sociology, Utrecht http://home.fsw.vu.nl/HBG.Ganzeboom/ISMF/index.htm
Lai, G., 2008. Marriage, gender and social capital. In: Lin, N., Bonnie, H. (Eds.), Social Capital: An International Research Program. Oxford University Press, Erickson, Oxford, pp. 342–363.
Lin, N., 1999. Social networks and status attainment. Annu. Rev. Sociol. 25, 467–487.
Lin, N., 2001. Building a network theory of social capital. In: Lin, N., Cook, K., Burt, R.S. (Eds.), Social Capital: Theory and Research. Aldine de Gruyter, New York, NY, pp. 3–29.
Lin, N., 2002. Social Capital: A Theory of Social Structure and Action. Cambridge University Press, New York, NY.
Lin, N., Ao, D., 2008. The invisible hand of social capital: an exploratory study. In: Lin, N., Bonnie, H. (Eds.), Social Capital: An International Research Program. Oxford University Press, Erickson, Oxford, pp. 107–132.
Lin, N., Dumin, M., 1986. Access to occupations through social ties. Social Netw. 8, 365–385.
Lin, N., Fu, Y., Hsung, R.M., 2001. The position generator: measurement techniques for investigations of social capital. In: Lin, N., Cook, K., Burt, R.S. (Eds.), Social Capital: Theory and Research. Aldine de Gruyter, New York, NY, pp. 57–81.
Magee, Marc Porter, 2008. Civic participation and social capital. A social network analysis in two American counties. In: Lin, N., Erickson, B.H. (Eds.), Social Capital: An International Research Program. Oxford University Press, Oxford, pp. 308–327.
Moore, S., Daniel, M., Gauvin, L., Dubé, L., 2009. Not all social capital is good capital. Health Place 15, 1071–1077.
Mouw, T., 2003. Social capital and finding a job: do contacts matter? Am. Sociol. Rev 68, 868–898.
Solon, G., 1989. Biases in the estimation of intergenerational earnings correlations. Rev. Econ. Stat. 71, 172–174.
Statistics Sweden, 1982. Socioekonomisk indelning (SEI). Meddelanden i samordningsfrågor 1982:4.
Van der Gaag, M., Appelhof, G.J., Webber, M., 2012. Ambiguities in responses to the Position Generator. Sociologia E Politiche Sociali.
Van der Gaag, M., Snijders, T.A.B., Flap, H.D., 2008. Position Generator measures and their relationship to other Social Capital measures. In: Lin, N., Erickson, B. (Eds.), Social Capital: An International Research Program. , pp. 27–49.
Verhaeghe, P.-P., Van de Putte, B., Roose, H., 2012. Reliability of Position Generator Measures across Different Occupational Lists: a Parallel-Test Experiment. University of Ghent.