THE SHADOW OF HIERARCHY - HOW TO SAMPLE A HIDDEN POPULATION OF FORMER EMPLOYEES?

Danny Pająk
European-University Viadrina, Frankfurt (Oder), Germany

Abstract

Relocations in terms of outsourcing to a non-affiliated company and offshoring, the cross-border relocation within the company, are widely used in recent years and in many cases cause collective employee layoffs. Even if one of the main intentions is the reduction of costs, relocations may not produce the highly anticipated financial benefits that most companies pursue. One reason is that organizations often have overlooked and underestimated social or 'hidden' consequences of relocations. The goal of the project was to investigate the research question whether there is a kind of hierarchical 'shadow'. Do former hierarchical structures still exist among victims and survivors of the relocation? How is this structure affected by hierarchy even years after the event and how is the shadow affecting the hierarchy of the firm itself? To answer these questions and to test whether former employees are still connected among one other, a pilot study was carried out among a German manufacturer of electrical equipment which relocated its entire workforce in 2006. The pilot study also tested the feasibility of Respondent-Driven-Sampling (RDS) as an effective and efficient form to sample rare and hidden populations. The paper will introduce the RDS methodology to the institutional research community and gives advice for the implementation of a RDS-study.

Keywords *Shadow of Hierarchy, Relocation, Offshoring, Former Employees, Social Network, Hidden Population, Respondent-Driven-Sampling (RDS)*

Correspondence: Danny Pająk, Chair of International Management, European-University Viadrina, Große Scharrnstr. 59, 15230 Frankfurt (Oder), Germany (E-Mail: pajak@europa-uni.de)

1. Introduction

Relocations in terms of outsourcing to a non-affiliated company and offshoring, the cross-border relocation within the company, are widely used in recent years and in many cases cause collective employee layoffs. Even if one of the main intentions is the reduction of costs, relocations may not produce the highly anticipated financial benefits that most companies pursue. One reason is that organizations often have overlooked and underestimated social or 'hidden' consequences of relocations.

While there is considerable research in general downsizing, *"the act of eliminating employees by permanent layoffs, cutbacks, attrition, early retirement and termination"* (Gandolfi, 2006, p. 73), relatively little empirical research has focused upon those people who are directly affected from relocations. This turns out to be a research gap since the laid off workforce could play a crucial role in the future success of firms. A recent study by Kinkel et al. (2007) reports that among every third to fourth German relocation at least one company relocates its production back within the first five years. Those back-shoring companies often re-employ former employees to save transaction costs in selecting, recruiting and training their workforce. For this reason these companies should care about the 'victims' of relocation decisions. However, the laid-off workforce is often neglected during the relocation process and far too little attention has been paid to questions like the willingness of re-employment or organizational endorsement.

The main reason for this research gap seems to be a lack of a known sampling frame and the rare prevalence of the target population among the total labor-force. According to a study of the German Federal Statistical Office, 188,600 workplaces in Germany have been relocated abroad between the years 2001 and 2006 (Destatis, 2008). Assuming that in 2006 the total German labor force consists of 43.3 million people, the target sub-population of relocation 'victims' counts only 0.44% in the parent population. According to Gabor (2007) subsets of general populations that generally represent less than 2% are defined as hidden, rare or hard-to-reach populations. They are characterized by the difficulty to use standard probability sampling methods, due to a lack of a known sampling frame and its small or rare member size. For a hidden population, like employees affected from relocations, a known sampling frame simply does not exist. Even if a frame could be constructed, a broad random sample of the overall population would not be cost-effective given the low base rate of relocation victims among the total German work-force. For instance, to interview one percent of the target population using traditional methods like household survey or random digit dialing, 430,000 people would have to be randomly sampled, which would be prohibitively expensive and infeasible.

For the specific type of hidden or rare populations several sampling methods like time, space, institutional or targeted sampling exist, which are all labor, time and resource intensive. To overcome some of these difficulties respondent-driven-sampling (RDS) was developed by Heckathorn (1997; 2002) as a refinement to traditional chain-referral or snowball sampling methods. RDS belongs to a family of sampling methods termed link-tracing or adaptive sam-

pling designs *"that are designed to be used in settings where traditional probability sampling methods are infeasible"* (Coryn, Gugiu, Davidson & Schröter, 2007, p. 5). For more than a decade RDS has been used now, primarily for AIDS-prevention projects. A recent review by Johnston et al. (2008) reports 128 RDS studies worldwide mainly including hidden populations of infection drug users, men who have sex with men and commercial sex workers.

In contrast to snowball samples (Goodman, 1961), where valid statistical inferences are not possible due to its convenience non-probability sampling approach and an unknown sampling frame of the hidden target population, RDS was primary designed to estimate population characteristics and prevalences within hard-to-reach populations. Although by using the RDS methodology one also starts with a convenience non-random sample, by identifying initial respondents reachable or identifiable for the researcher, it is assumed that the sample composition reaches an equilibrium within a limited number of recruitment waves. This means that after reaching the equilibrium all social ties of the target population have an equal probability of being recruited (Salganik & Heckathorn, 2004). The inclusion probabilities of respondents into the sample will be weight to reduce or eliminate biases resulting from the choice of initial subjects (Heckathorn & Jeffri, 2003). According to Gile and Handcock (2010) this is the main advantage of RDS over non-probability sampling methods like traditional snowball sampling approaches. The current RDS estimator used for calculating inferences is based on a first-order Markov Chain assumption. Further theoretical principles of this underlying stochastic procedure are beyond the scope of this paper (Goel & Salganik, 2009; Heckathorn, 1997; Heckathorn, 2007; Heckathorn, Semaan, Broadhead & Hughes, 2002; Salganik & Heckathorn, 2004; Volz & Heckathorn, 2008).

The objective of this article is twofold. The main objective is to test whether RDS is an appropriate method to sample social networks of former colleagues that were relocated or offshored. A second goal of this article is to show that even years after a termination process a kind of hierarchical shadow exists between 'victims' and 'survivors' of a relocation decision. This finding has some important implications for companies trying to reduce their costs through relocations. How is this 'shadow' affected by hierarchy even years after the event and how does it affect hierarchy? In a recent review about the close relation of organization theory and Oliver Williamson's transaction cost theory, Baudry and Chassagnon (2010) emphasize the importance of organizational atmosphere as part of the informal organization. Williamson argues that *"supplying a satisfying exchange relation is part of the economic problem"* (Williamson, 1976, p. 372). Although it is widely accepted that organizational atmosphere plays a critical role in the settlement of disputes in organizations, it is *"rarely analyzed in the literature devoted to transaction cost economics"* (Baudry & Chassagnon, 2010, p. 491).

This paper proceeds in the following steps: In Section 2, the principles of Respondent-Driven-Sampling (RDS) and its assumptions are presented. Section 3 discusses the implementation issues and requirements that have to be fulfilled in formative research before using the RDS methodology. Section 4 contains information about the method used in the feasibility study. In Section 5, the results of the network analyses are presented and the effects of those

results for the usage of the RDS methodology are illustrated. Additional recommendations for the implementation are discussed. Section 6 concludes and outlines the limitations and suggested future directions necessary for a deeper understanding of the role hierarchical shadows play in implementing relocation decisions.

2. The principles of Respondent-Driven-Sampling (RDS)

In RDS participants recruit their peers, rather than identifying them to the researcher as is the case in traditional so-called snowball samplings (Johnston & Sabin, 2010, p. 39). The direct contact between recruiter and recruit based on a pre-existing reciprocal relationship is one of the key elements that distinguish RDS from other forms of chain-referral sampling (see Tiffany, 2006, p. 114). By using data on the personal network size of the respondent in the target population as well as information on the sociometric structure (e.g., gender or status) of its self-reported acquaintance network, RDS controls for several biases usually introduced by the non-random selection of the initial respondents (called 'seeds' in RDS) in traditional chain-referral samples. Additionally, RDS provides information for analyzing the social structure of the population through network analysis of the recruitment chains. Thus, Wejnert concludes that *"RDS utilizes the reach of snowball sampling while maintaining the ability to make unbiased statistical inference"* (Wejnert, 2010, p. 113).

Sources of bias in traditional chain-referral methods include (see Heckathorn, 2002a; Johnston, 2008; Wang et al., 2005): (1) the non-random selection of seeds, (2) volunteerism, which means that more cooperative respondents are more likely to enroll in the study, (3) masking, in which respondents protect potential recruits by not referring them, (4) differentials in recruitment, which means that some respondents are more willing to recruit other participants, (5) different social network size, which implies that people with many peers are able to recruit more participants and therefore oversampling those groups and (6) the tendency for in-group recruitment, where initial respondents tend to recruit participant with similar sociometric or demographic characteristics such as gender or status level, which is also know as tendency of homophily.

Long recruitment chains will reduce those biases induced by the convenience selection of seeds. Furthermore, they will insure a non-zero probability of inclusion into the recruitment procedure. Therefore, Heckathorn and colleagues (2002) lists three reason why long referral chains are desirable. First, the longer the recruitment chains, the greater the chance to reach the most hidden and isolated individuals of the population, who have little contact with other members of target population. Second, given the theoretical geometrically growth of the recruitment process, *"longer referral chains ensure that the sociometric distance between the seeds and the bulk of the sample will be large, thereby further assuring the diversity and representativeness of the sample"* (Heckathorn et al., 2002, p. 60). Finally, since the aim of RDS is to reach an equilibrium where the sample composition is independent from the initial non-random selection of seeds, long referral chain will ensure that the sample stabilizes, even if the equilibrium is reached comparatively slowly (see Heckathorn, 2002a, p. 60; Wang et al., 2005, p. 148). Therefore, as Coryn and colleagues (2007) highlight, the most important de-

sign element of RDS is the use of techniques to ensure long-lasting referral chains over several recruitment waves.

The techniques used in RDS to produce long referral chains are a unique-numbered recruitment coupons and a dual incentive system. In most RDS studies, each participant receives a fixed number of three coupons to recruit members of their own social network who are part of the target population (see Johnston et al., 2008, p. 295). Usually this includes the seeds as well as each additional respondent who is participating in the survey throughout the recruitment process. Limiting the coupons address the biases of differential recruitment and volunteerism introduced above. According to Johnston et al. *"the use of three coupons throughout RDS recruitment has a practical, rather than theoretical, rationale"* (Johnston et al., 2008, p. 297). The number of coupons distributed to the respondents should rather vary according to the wave reached in the recruitment process. By reducing the number of coupons (e.g. from three coupons to two coupons after the 5th wave) the sample will not only consist of longer recruitment chains but it will also be *"more closely suited to the theoretical principles of linear recruitment [...] upon which RDS is designed"* (Johnston et al., 2008, p. 297). The systematic coupon reduction process after a certain wave ensures that slower recruitment chains can catch-up. Thus the tendency for in-group recruitments (homophily biases) can be further reduced, which increases the chance for reaching the most hidden or isolated members of the target population (Johnston et al., 2008).

The second important technique used in RDS is the dual incentive system. First of all, respondents are usually rewarded with an incentive for participating in the survey and secondly, they also get additional rewards for each peer of their social network that they are able to convince to participate. The second incentive produces some kind of peer-pressure which will reduce the volunteerism bias as well as further increase the chance for long recruitment chains (Wang, Falck, Li, Rahman & Carlson, 2007). According to Magnani et al., the second incentive *"can* [also] *be expected to reduce non-response bias because those who would not participate for financial reasons alone may do so as a favor to a friend"* (Magnani, Sabin, Saidel & Heckathorn, 2005, p. 71). The combination of coupon distribution and rewarding will also produce some kind of social pressure. The more coupons somebody receives from different recruiters over several waves during the recruitment process the higher the pressure and chance that this person will also participate in the survey. Wejnert and Heckathorn (2008) named this kind of effect a 'norm of participation'. Another important advantage of the dual incentive system is, that the recruitment can be guided by the researcher. By using steering incentives as additional bonuses one can reach predefined specific subgroups inside the target population (Wang et al., 2005; Wejnert, 2010). All given incentives must be appropriate to the studied population. Johnston and colleagues suggests that they should not be too high or too low, *"to motivate economically and socially diverse members of the target population"* (Johnston et al., 2008, p. 295). Especially the secondary incentive should be only moderate. Otherwise the peer pressure could become coercive (Heckathorn et al., 2002) and people who are not part of the target population attempt to participate to get the financial rewards (Salganik & Heckathorn, 2004). Besides monetary compensations, other incentives

like lotteries, vouchers or donations to charities are possible, depending on the financial resources of the researcher and the characteristics of the hidden population.

Since RDS uses each personal network size to calculate the probability of the selection, three additional pieces of information have to be collected in a study using the RDS-methodology: (1) the size or degree of each respondents personal network in the target population (e.g., 'How many former colleagues, friends or relatives do you know who were also affected from a relocation in the last five years and with whom you had personal contact during the last year?', (2) the structure of those self-reported network (e.g., gender by asking 'How many of those people you had contact with in the last year are male?') and (3) information about the pre-existing social relationship (e.g. Who gave you the coupon to participate in this study: a former colleague, a close friend, a relative, an acquaintance or a stranger) (Johnston & Sabin, 2010; Wang et al., 2005). Additionally, all distributed coupons are also tracked through their unique serial number, which allows the researcher to link each recruit with its recruiter (see Wang et al., 2005, p. 148).

While the original RDS estimation process was characterized by a set of six assumptions (Salganik & Heckathorn, 2004) a recent work by Heckathorn (2007) reduces the set to five underlying theoretical assumptions (see Wejnert & Heckathorn, 2008, p. 112):

(1) *Respondents maintain reciprocal relationships with individuals whom they know to be members of the target population;*

(2) *Respondents are all linked to a single component in the network;*

(3) *Sampling is with replacement;*

(4) *Respondents can accurately report their personal network size or, equivalently, their degree;*

(5) *Peer recruitment is a random selection of the recruiter's peers.*

There are a lot of controversial discussions about these theoretical assumptions. Especially the last three are discussed in a lot of detail that are beyond the scope of this paper and can be found elsewhere (Gile & Handcock, 2010; Heckathorn, 2002b; Volz & Heckathorn, 2008; Wejnert, 2009; Wejnert & Heckathorn, 2008).

The five assumptions are necessary for the sample to stabilize and to reach equilibrium, which means that the *"sample distribution on key variables* [e.g., gender or status] *remains stable within 2% of the equilibrium distribution, even though more individuals enter into the sample"* (Johnston & Sabin, 2010, p. 40). This means, as the sample grows from wave to wave the composition of key variables ceases to change (see Heckathorn et al., 2002, p. 57). Assuming that the true population value of blue-collar workers in the target population is 58 percent, equilibrium would mean that the sample composition will stabilize around this value after a certain number of recruitment waves. In other words, the ratio between blue- and white-collar workers would remain stable (within 2%), no matter how many additional individuals are recruited into the sample (see Johnston, 2008, p. 15). In theory, the equilibrium

will be usually reached within four to six waves (Magnani et al., 2005; Wejnert, 2010). After the wave where the equilibrium is reached it can be assumed that the sample becomes representative for the target population and thus independent from the non-random selection of the initial respondents (Johnston & Sabin, 2010). Diversifying the initial respondents will help to reach the equilibrium in a manageable number of waves. The ideal seed is often described as a sociometric star, well motivated and enthusiastic (see Wejnert & Heckathorn, 2008, p. 109).

But besides diversifying the seed selection, paying attention to the five assumptions and gathering additional required information from the respondents (e.g., personal network size), RDS is nothing more than a good chain-referral sample, unless analyzed with a specialized software (see Johnston, 2008, p. 302). The most common used software is RDSAT, which can be downloaded from http://www.respondentdrivensampling.org in the stable 5.6 version (Volz, Wejnert, Degani & Heckathorn, 2007).

3. Implementation Issues and Requirements for using RDS

Although RDS was primarily developed for AIDS-prevention projects, there are examples where the methodology was used for sampling different populations. Heckathorn and Jeffri (2001; 2003) for example used RDS to sample Jazz musicians in New York City and San Francisco. Usually, storefront interviews in locations accessible to the target population are used to implement the survey. But, there have also been first attempts to deviate the sampling process from face-to-face interviews. Hagan (2001) uses RDS in telephone interviews to study American Vietnam War draft resisters in Canada (Salganik & Heckathorn, 2004). The methodology was also applied among social networks of students by using e-mail referrals (Wejnert, 2010; Wejnert & Heckathorn, 2008).

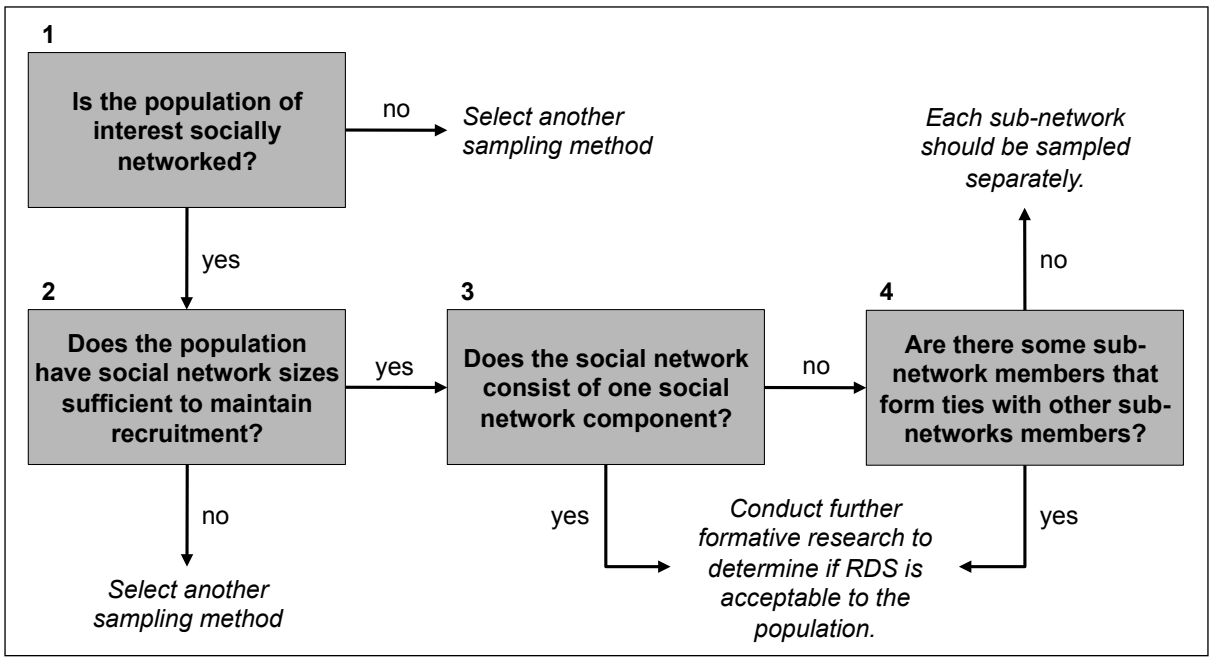
However, despite these examples, the target population in this study seems to be completely different. To the author's knowledge it is the first time that the feasibility for using the RDS methodology is tested among a population where the referral-chains rely on a former network of interactions. Since the network and its former hierarchal structures do not longer exist, one aim of the feasibility study was to test whether there is a kind of hierarchal shadow among victims and survivors of relocations where communication and interaction is still taking place.

Former employee networks are the source of some potential limitations for using RDS. The main problem is that the former network is shrinking, which could result in a disconnected network. Depending on the time-frame between the intervention into the network structure by laying off employees and the time when the recruitment of those former colleagues takes place, the size of the reachable network may varies. Nodes within the network may get isolated because of the removal of ties or links between former colleagues. There are several potential sources for the shrinking effect in those kinds of networks. One problem is the increasing average age of the population the longer the time-frame between the relocation and the sampling time gets. Former ties between colleagues can be eliminated due to illness and death of former colleagues. Another danger lies in the job-seeking behavior of the laid-off

employees. One can assume the larger the distance between the former workplace and living environment and ones new center of life gets, the less frequent the interaction with former colleagues will be. A third group of problems exists because of changing living conditions. Telephone numbers, address information or even the surname (i.e., divorce or marriage) can change over time, which make it difficult to reach or contact those former colleagues with a chain-referral approach.

Therefore according to Heckathorn and Jeffri (2005) four requirements must be met when someone wants to use RDS to sample a hidden population: (1) Recruiter and recruit must know another as members of the same target population, (2) the network must be dense enough, (3) the population should not be so highly clustered into several sub-groups and (4) *"means must exist for motivating respondents to recruit their peers"* (Heckathorn & Jeffri, 2005, p. 3). According to Johnston and Sabin (2010) the questions from Figure 1 should therefore be answered first before using RDS for sampling hidden populations. If one of the first two questions is answered with 'no', another sampling method should be selected. But even if question three or four are affirmed, further formative research should be conducted to determine whether RDS is acceptable to the population.

Figure 1: How to decide whether RDS is a suitable sampling method?



Source: Own work based on a table presented by (Johnston & Sabin, 2010, p. 45)

The questions of Figure 1 will be answered in Section 5. In order for RDS to work, the target population needs to form a single connected network with a fairly high frequency of interaction among former colleagues.

4. Methods and Data Analyses

The author used a manufacturer of electrical equipment which relocated its entire workforce from Berlin to a rural area in the federal state of Lower Saxony in 2006 to test the feasibility

of RDS to sample victims of relocations. Approximately three and a half year after the relocation had been realized the survey was carried out from the end of December 2009 to March 2010. In three qualitative interviews, 96 former employees of that company could be identified. This number is similar to press releases about the relocation at that time, where 95 to 100 employees affected from the relocation were reported. 90 of those identified were eligible for the study purpose. Five former employees were excluded from further analysis either because they were laid-off before the relocation or they emigrated back to their native country in the meantime. Another person exceeded the 65-year limit, which is above the limit to study the willingness of reemployment in a following survey. From 37 former colleagues, none or wrong contact information could be found in the three preliminary interviews. This means only 53 questionnaires could be distributed among the former company members, which were filled out by 28 interviewees giving a response rate of 53 percent (for details see table 1).

Table 1: Overview of the Sample Population

	n	in %	Female	Male	Blue collar	White collar	Victim	Survivor
Overall identified population	**96**	**100%**	**57 / 59.4%**	**39 / 40.6%**	**57 / 59.4%**	**39 / 40.6%**	**72 / 75.0%**	**24 / 25.0%**
Not included in the analysis	6	6.3%	4 / 4.2%	2 / 2.1%	5 / 5.2%	1 / 1.0%	5 / 5.2%	1 / 1.0%
- Emigrated in the meantime	3	3.1%	1 / 1.0%	2 / 2.1%	2 / 2.1%	1 / 1.0%	2 / 2.1%	1 / 1.0%
- Over 65 years in the meantime	1	1.0%	1 / 1.0%	-	1 / 1.0%	-	1 / 1.0%	-
- Laid-off before the relocation	2	2.1%	2 / 2.1%	-	2 / 2.1%	-	*2 / 2.1%	-
Initial Sample (93.8% of 96)	**90**	**100%**	**53 / 58.9%**	**37 / 41.1%**	**52 / 57.8%**	**38 / 42.2%**	**69 / 76.7%**	**21 / 23.3%**
Address-related losses	37	41.1%	30 / 33.3%	7 / 7.8%	30 / 33.3%	7 / 7.8%	30 / 33.3%	7 / 7.8%
- No contact information	25	27.8%	23 / 25.6%	2 / 2.2%	25 / 27.8%	-	23 / 25.6%	2 / 2.2%
- Undelivered returned to sender	9	10.0%	5 / 5.6%	4 / 4.4%	5 / 5.6%	4 / 4.4%	6 / 6.7%	3 / 3.3%
- E-Mail address bounced	3	3.3%	2 / 2.2%	1 / 1.1%	-	3 / 3.3%	1 / 1.1%	2 / 2.2%
Eligible Sample (58.9% of 90)	**53**	**100%**	**23 / 43.4%**	**30 / 56.6%**	**22 / 41.5%**	**31 / 58.5%**	**39 / 73.6%**	**14 / 26.4%**
Non-Response	24	45.3%	9 / 17.0%	15 / 28.3%	7 / 13.2%	17 / 32.1%	14 / 26.4%	10 / 18.9%
Refused	1	1.9%	-	1 / 1.9%	1 / 1.9%	-	1 / 1.9%	-
Survey responses	**28**	**52.8%**	**14 / 26.4%**	**14 / 26.4%**	**14 / 26.4%**	**14 / 26.4%**	**24 / 45.3%**	**4 / 7.5%**

* These two were laid-off before the decision to relocate the company site. This means they are not victims of the relocation.

The primary goal of the questionnaire was to identify the social network structure among those former employees. Besides a general questionnaire, each interviewee received a list of all the identified names and they were asked to which of those former colleagues they still had contact in 2009. Although only the self-reported out-degree of the individual respondents was measured in the study (number of contacts to other colleagues in the year 2009), under the reciprocity assumption of RDS it is assumed that the out-degree equals the in-degree (Wejnert, 2010). Thus, unless otherwise stated, the results in Section 5 correspond to an undirected analyses of the identified network, presuming that when respondent A reported a contact with respondent B, B would also report a contact with A.

5. Results and Discussion

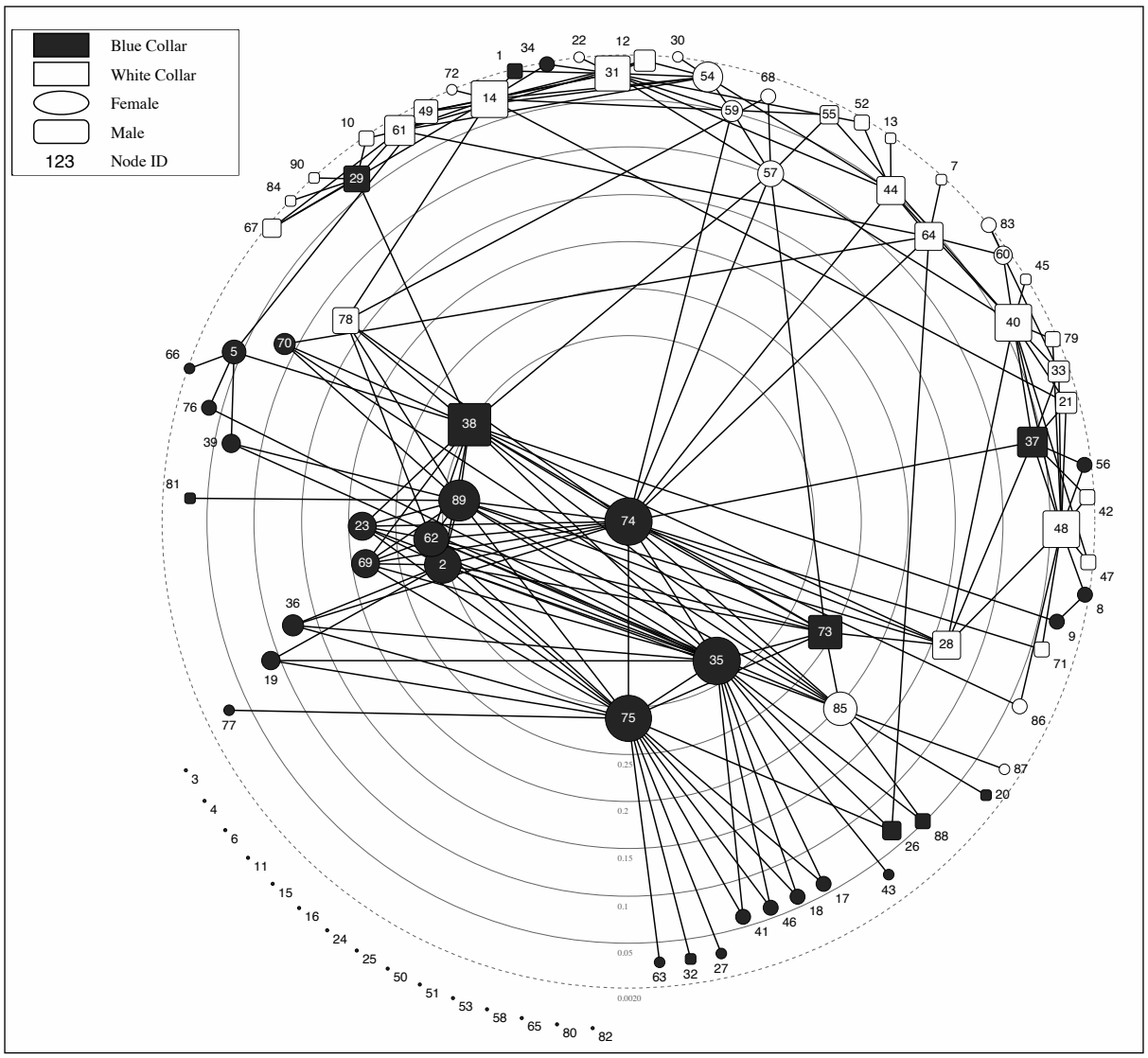
The Giant Component and Isolated Sub-Groups

There have to be at least five links on average from one node to another to ensure long referral chains (Heckathorn, 1997; Heckathorn et al., 2002; Wejnert & Heckathorn, 2008). Apart from long referral chains, this will ensure that the network will consist of one social network

component (see the third question in Figure 1). As Wejnert et al. state: *"in a random network, a single component forms when individual degrees* [i.e., the amount of links to former colleagues] *are large compared to the natural log of the population size"* (Wejnert & Heckathorn, 2008, p. 112). Volz et al. (2008) point out that most social networks posses a so-called giant component, which means that every person is connected with every other person in the network. In pure random graphs the giant component will consists of 99% of the population if nodes have just five links on average (Volz & Heckathorn, 2008, p. 83).

The average degree of the respondents in Figure 2 was 8.6 which is above the threshold. Successful studies using RDS have also been carried out with values below that degree. Heckathorn and Jeffri (2005) for example list a RDS survey that was carried out among Latino men who have sex with men in San Francisco which had a self-reported average network size of only six degrees (see also Ramirez-Valles, Heckathorn, Vázquez, Diaz & Campbell, 2005).

Figure 2: Network Displayed in Eigenvector Centrality Layout (Node Size = Degree)



Source: Own work analyzed with Visone (http://www.visone.info).

Although only the answers of 25 respondents (28% of the initial sample size of n=90) were usable for the network analysis, the giant component already contains 75 (83.3%) former colleagues (see Figure 3). Consequently the number of people identifiable through the network approach exceeds the number of eligible respondents identified through the three preliminary qualitative interviews. Figures 2 and 3 also show that there are bridges between people of various sub-groups in the network (i.e. between blue and white collar workers and between victims and survivors of the relocation). But what Figure 2 also shows is that there are still 15 isolated nodes in the network from which no path to another node in the component exists. While all 38 former white collar workers are part of the giant component (see also Table 1), 15 former blue collar workers are still missing. 14 of them are female with a high proportion of members from ethnic minorities (i.e., 6 of them are Turkish, 4 are Yugoslav, 1 is Polish and only 3 are German). It seems therefore extraordinarily difficult to sample those sub-population of female blue collar workers who are immigrants. One possible solution to get access to this kind of sub-group could be to provide steering incentives as mentioned in section 2 of this paper.

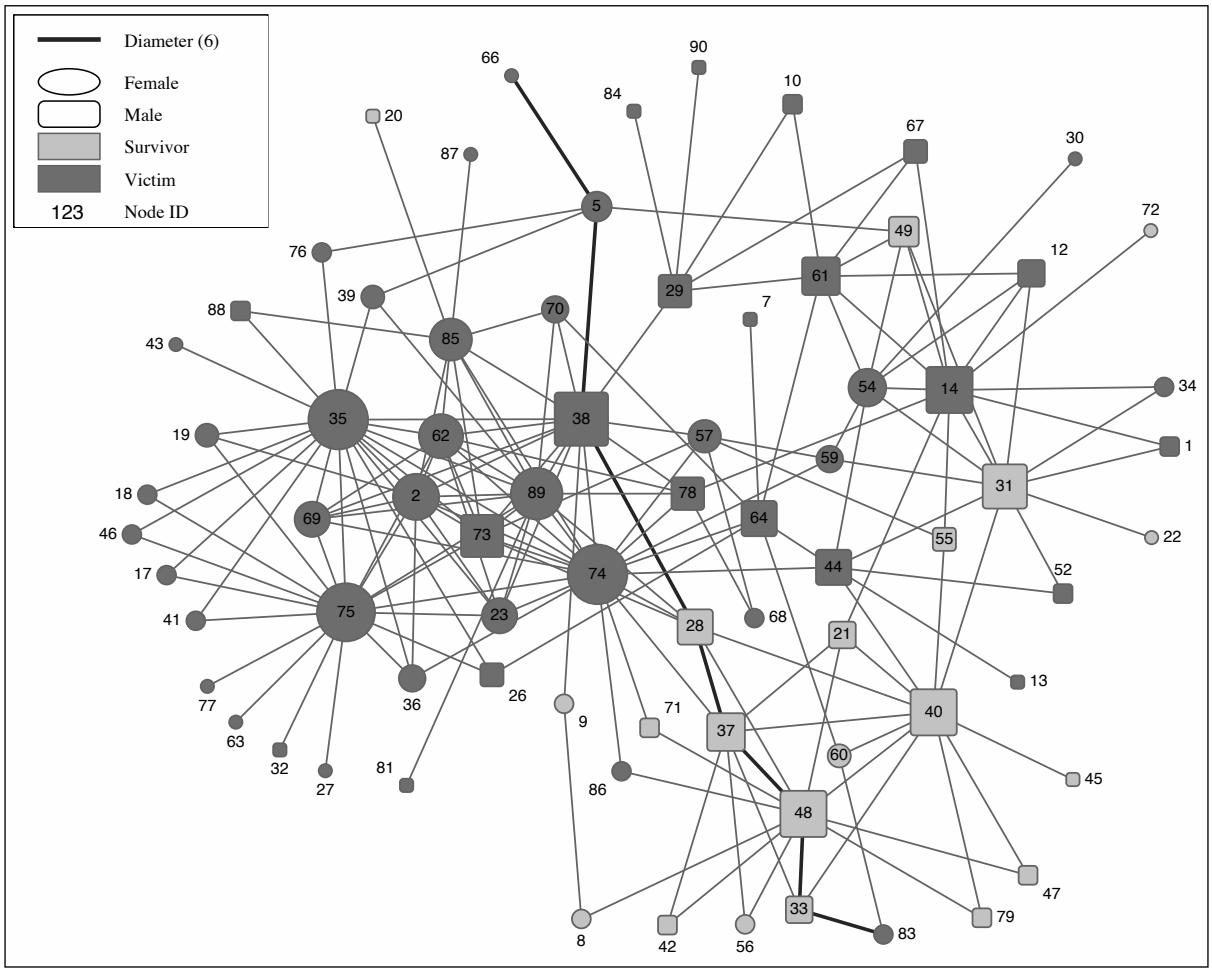
Another approach was used by Heckathorn and Jeffri (2003). Since they were concerned in their study about Jazz musicians in New York and San Francisco that not enough female musicians could be recruited, they implemented a coupon quota. *"Three coupons could be given to any jazz musician, but one had to be given only to a female jazz musician"* (Heckathorn & Jeffri, 2003, p. 51). Gile (2010) recommends to use additional information whenever possible to verify the relative size of sub-populations. Sometimes for example, the gender composition of the relocating company can be found in press releases. Other source like annual reports, unions, employment agencies or transfer and transitional companies may be used to get further insights into the population composition, which can be used to implement coupon quotas. When the resources are available it can further help to translate the questionnaire into different languages. This will reduce the masking bias due to beliefs that the questions would be to difficult to understand for immigrants, which could hinder the recruitment of those sub-groups.

The Network's Diameter

Figure 3 also shows the diameter of the giant component, which is the longest shortest path of the network. The diameter of the network is 6, corresponding to the distance between nodes 66 and 83. Assuming that node 66 would be taken as a seed, at least six referral chains would be necessary to reach the victim with ID number 83. The diameter of a network is closely related to works on the so-called small-world phenomenon. By using a letter-referral technique, the social psychologist Stanley Milgram (1963) found out that the average path lengths between any pair of people in a population is quite short (Schnettler, 2009b). Thus, as Salganik and Heckathorn (2004) state, one *"can be fairly confident that with reasonably long recruitment chains all members of the population have a nonzero probability of being reached even if the seeds occupy an unusual place in the network structure".* Although later studies have found an even shorter average path length, Milgram's work was the basis for the later usage of the 'six degrees of separation' rule in everyday language (see Schnettler, 2009a for a

comparison of different small-world studies).

Figure 3: Giant Component with Network Diameter

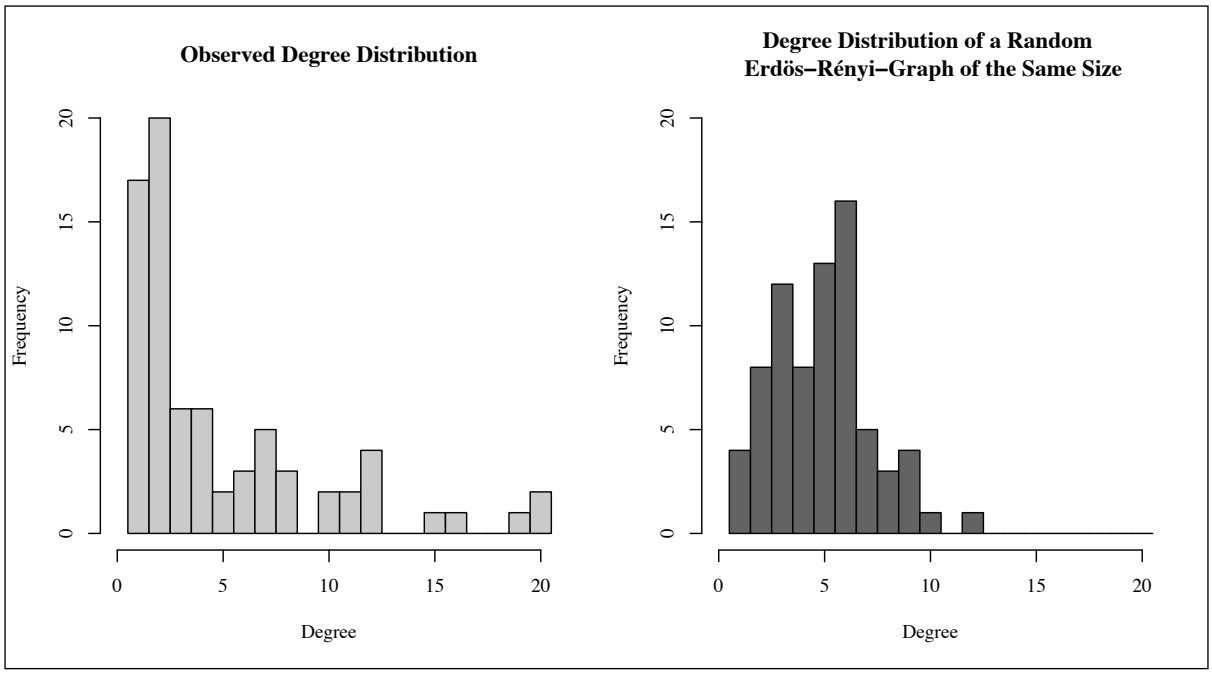


Source: Own work analyzed with Visone (http://www.visone.info).

The Degree Distribution of the Network

A third easily identifiable property of the network is its degree distribution. While the degree distribution in a random graph follows a bell curve, social networks often show a power law distribution (Jackson, 2006). That means, power-law or scale-free networks are characterized by a few highly connected nodes and many nodes with only a few connections. Most social networks are scale-free. Figure 4 shows a comparison of the observed degree distribution in the feasibility study compared to a degree distribution of random network with the same size. Besides a high amount of people that have only one or two connections there are also a few nodes with up to 20 links to former colleagues. Those highly connected nodes are known as 'hubs' in network analysis. Hubs play a key role in keeping the network together (Ravasz & Barabási, 2003), which is especially important for the robustness of a shrinking network like those of former colleagues. Additionally, hubs are also important for bridging different sub-populations in the network. For example, a hub node that functions as a bridge between victims and survivors of the relocation (e.g., the node with ID 38 in Figure 3) will have important implications for the spread of information or rumors between those two sub-groups.

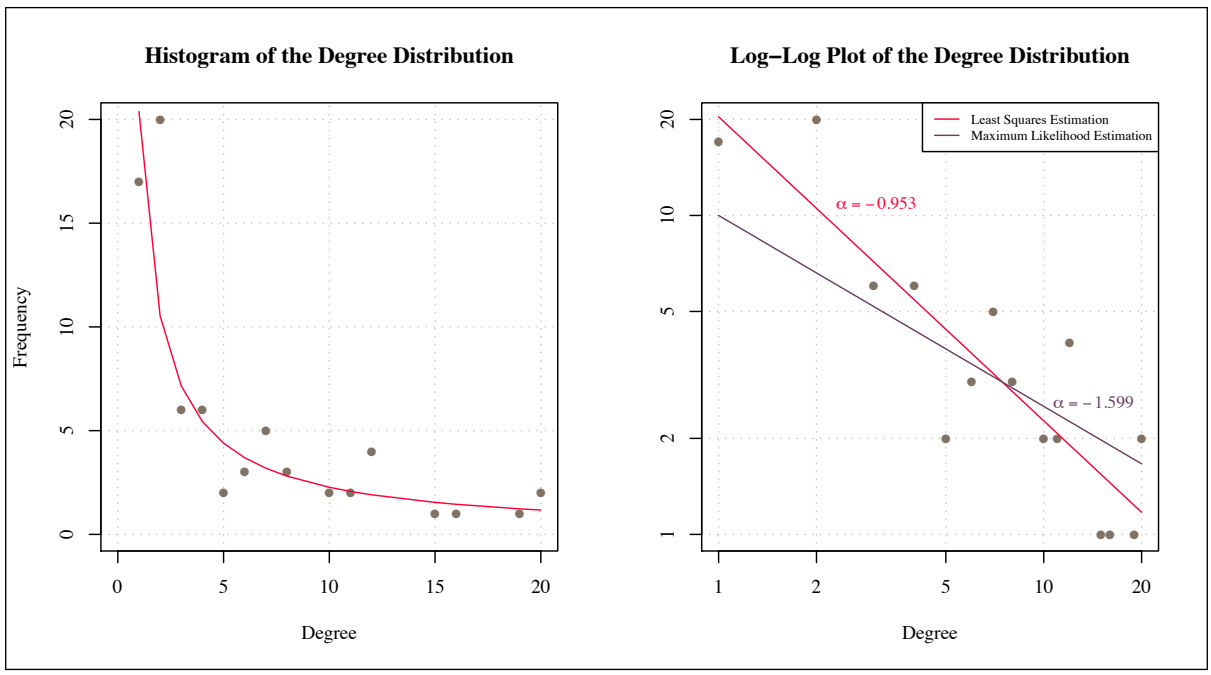
Figure 4: Observed Degree Distribution versus the Degree Distribution of a Random Graph



Source: Own work created with R.

A standard approach to outline the degree distribution of a social network is to plot the histogram on a logarithmic scale and see whether this looks linear (Jackson, 2006; Newman, 2005). Figure 5 illustrates the power law relationship of the studied network.

Figure 5: Degree Distribution and Log-Log-Plot



Source: Own work created with R. The right panel of Figure 5 shows the power-law distribution ($p(x) = C^{x-\alpha}$) plotted on a logarithmic scale, where the constant α is called the exponent of the power-law. The exponent value was calculated in two ways: (1) by fitting a straight line through the data points and (2) by using the maximum likelihood method described in further detail by Newman (2005).

Réka et al. (2000) demonstrated a high degree of robustness for scale-free networks. They show that even when as many as 5% of the nodes in a scale-free network are eliminated, *"the communication between the remaining nodes in the network is unaffected"* (Réka et al., 2000, p. 380). Also the length of the network's diameter remains unchanged. Therefore, the identification of a power-law distribution has important implications for our shrinking network expectation. The removal of unimportant nodes (those with just a few connections) will not influence the overall network structure. On the other hand, scale free networks have a high vulnerability to the removal of highly connected hubs in the network, for instance through death of a person with a high amount of links to other network members. The network's diameter would increase and the structure of the network could break into many isolated clusters (Réka et al., 2000).

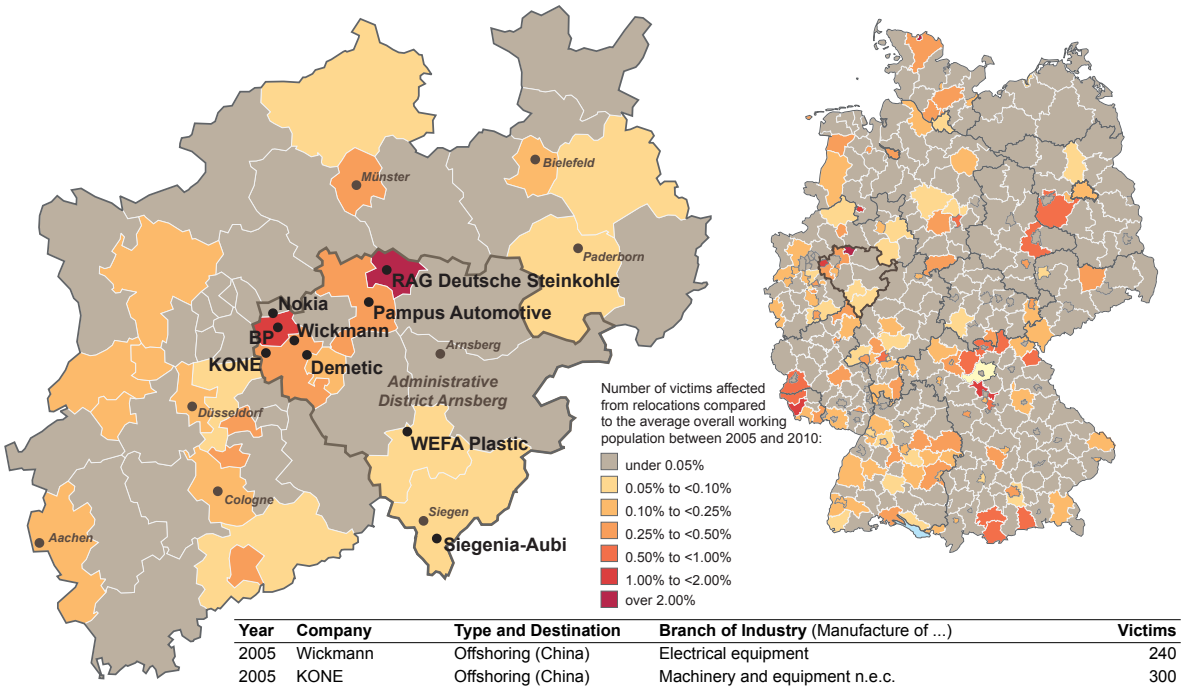
Seed Selection Process

Statistics about relocations in Germany are based extensively on estimations, which makes it difficult to define and locate the target population. That is why the author started to build an own database on German relocations. The German Relocation Database[1] (G.R.D.) collects news about company relocations from the trade, business and daily press. At the moment the G.R.D. contains detailed information about approx. 200 relocations between the years 2005 and 2010. Figure 6 shows a choropleth map of the German counties and autonomous municipalities on the right side with these 200 relocations. The different colors reflect the number of victims affected from relocations compared to the average overall working population[2] in the respective district between 2005 and 2009. The darker the various shades of red the more victims of relocations can be found in this area. The left side of Figure 6 shows a enlarged representation of the federal state of North Rhine-Westphalia with one of its administrative districts Arnsberg, where nine recorded relocations occurred in the investigated period. The table below the two choropleth maps lists detailed information about these nine companies (i.e., the year of announcement, the type and destination of the relocation, the branch of industry as well as the number of victims according to press releases).

What are the implications for the study purpose? First of all, there are different clusters in Germany where a comparatively high number of people affected from relocations can be found. The average working population in the administrative district of Arnsberg between 2005 and 2009 was approx. 1.7 million. By summing up the number of victims in Figure 6 one gets approx. 5,400 people that make up 0.32% of the working population in the district of Arnsberg, which means among every thousandth are at least three people affected from a relocation. Thus, besides high clustering in certain areas, the target population is rarely distributed among the total workforce. According to the data from the G.R.D. there are only two counties where the boundary value is above 2%, which is the generally the threshold value under which the target population is defined as a hidden or rare population (Gabor, 2007).

1. The database can be found under: http://www.ima-research.eu/projekt/verlagerung/datenbank/
2. For county level data see: German Federal Statistical Office and statistical Offices of the Länder, 2010

Figure 6: Company Relocations in Germany between 2005 and 2010



Year	Company	Type and Destination	Branch of Industry (Manufacture of ...)	Victims
2005	Wickmann	Offshoring (China)	Electrical equipment	240
2005	KONE	Offshoring (China)	Machinery and equipment n.e.c.	300
2007	Nokia	Nearshoring (Romania)	Computer, electronic and optical products	2,300
2007	Demetic	Domestic Relocation	Machinery and equipment n.e.c.	160
2008	WEFA Plastic	Domestic Relocation	Rubber and plastic products	60
2009	BP	Nearshoring (Hungary)	Coke and refined petroleum products	200
2009	Siegenia-Aubi	Nearshoring (Poland)	Fabricated metal products, except machinery and equipment	130
2010	RAG Steinkohle	Domestic Relocation	Mining of coal and lignite	1,622
2010	Pampus Automotive	Nearshoring (Czech Rep.)	Motor vehicles, trailers and semi-trailers	380

Map Source: Own work based on Wikimedia Commons, CC-BY-SA 3.0 Landkreise, Kreise und kreisfreie Städte in Deutschland.svg by TUBS

Source: Own work based on data of the German Relocation Database (G.R.D.).

Second, since the underlying overall network is disconnected in several firm or regional clusters or sub-populations, the representation of the target population will be determined by the seed selection (see Gile & Handcock, 2010, p. 322). Although in clusters like in the northwest of the administrative district Arnsberg, several bridges between sub-populations (e.g., a friendship or relationship between one person that worked for Nokia in Bochum and another relocation victim that work for a different company) are theoretically possible, it can be assumed that those ties between different sub-population members (i.e., victims of different firms) are relatively rare. Goel and Salganik use the term 'bottleneck' *"to refer to features of the network that lead to a propensity for the sampling procedure to get 'stuck' in certain regions"* (Goel & Salganik, 2009, p. 2208). Thus, the bias introduced by the non-random selection of seeds cannot be eliminated by simply increasing the length of the recruitment chains (see Gile & Handcock, 2010, p. 322). Therefore, in populations with bottlenecks between sub-groups the seed selection becomes increasingly important, so that the RDS methodology should be used separately on each sub-group (Gile & Handcock, 2010; Goel & Salganik, 2009).

As already stated in the last section, hubs play in important role in network analysis. To put it another way, those hubs are the sociometric stars one is looking for in the selection of initial respondents or seeds in RDS. They function as bridges between different sub-groups or sub-populations, have a high visibility and reputation among the target population and are there-

fore crucial for the recruitment process. How can those sociometric stars among people affected from relocations be identified and recruited? In newspaper articles about a relocation one can often find the names of employee representatives (e.g., work council members) interviewed in the course of the relocation or termination process. Those people can than be identified for instance by searching them in telephone directories. Another approach also used by the author for a larger research project about hidden costs of company relocations scheduled for 2011, is the identification of relocation victims through social network platforms in the internet. More precisely, the network by the author to find the seeds for that study is called Xing. Xing is the largest German social network to establish and maintain especially business contacts and relationships. According to a recent study of the German Federation of the Information Industry, Telecommunications and New Media (BITKOM), Xing reaches nine percent of the Internet users in Germany (Berg, 2011).

Combing the knowledge from the G.R.D. (i.e., year of the relocation, region and name of the relocating company) with Xing, one can search the employment history of members (see Appendix 1). The extended search option in Xing allows to search for previous employers in a certain region. For instance, by using the search terms *44** and *Nokia* one would find all Xing members that previously worked for Nokia in the area where the German postal code starts with 44 (i.e., Bochum). Since we know from the G.R.D. that the relocation took place in 2008, we can identify those people whose occupation at Nokia ended around 2008. Those can than be recruited as seeds for RDS. Wang et al. (Wang et al., 2005, p. 148) conclude that *"seeds do not necessarily need to be representative of the target population"*. Thus, survivors of the relocation could be used as seeds too, as long as they have contact to former colleagues that have been directly affected from layoffs.

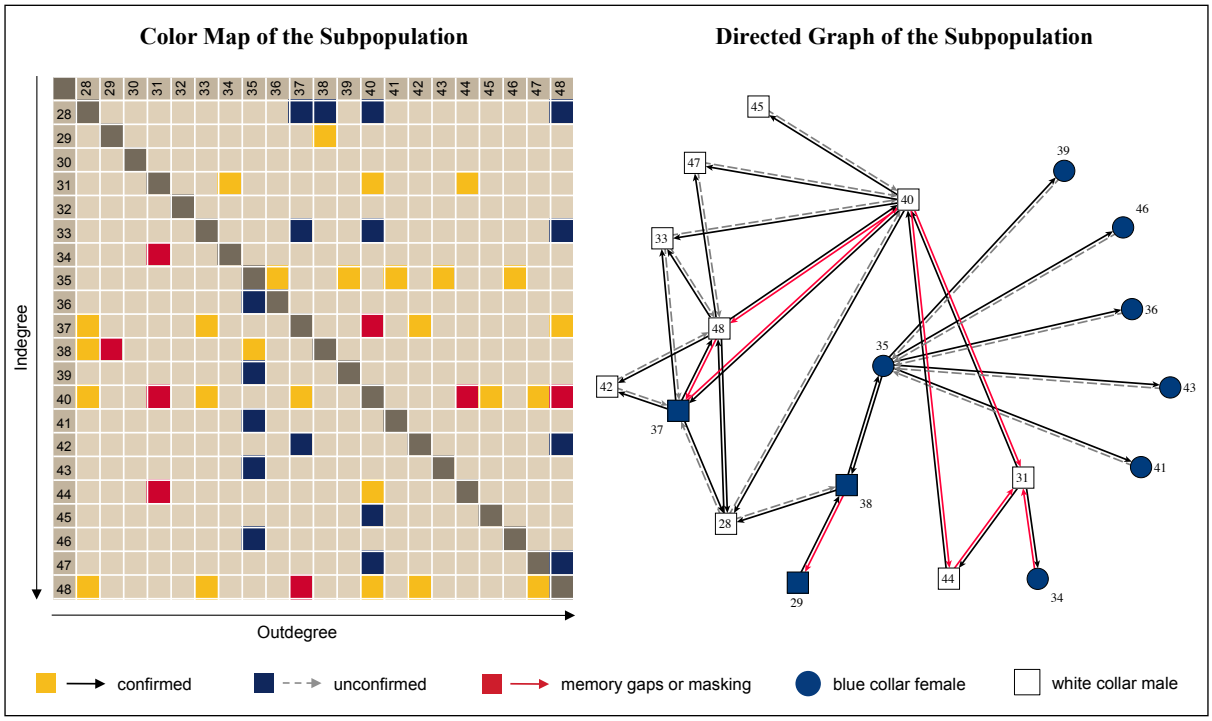
Memory Gaps in Degree Reporting

One important assumption of RDS is, that respondents are able to accurately report their personal network size. For Wang and colleagues (2007) this is one important limitation of RDS. They argue that respondents may not be able to accurately report their degree or number of contacts they had with members of the target population in a certain timeframe, an issue also found in this study. Figure 7 shows in the left panel a color map of a sub-population (i.e., nodes from ID 28 to 48) of the network. The contact patterns are differentiated in three ways: (1) a confirmed relationship from both nodes, (2) an unconfirmed relationship (i.e., only the sender reported the relationship and the confirmation of the receiver is missing due to non-response of that node) and (3) forgetting or masking of a relationship. The last mentioned contact pattern implies that for instance the person with ID 29 reported a relationship to node 38. Looking the other way around the person with ID 38 forgot to report this relationship with node 29, either because of simply forgetting to report it (memory gap) or because of masking that relationship. The right panel of Figure 7 shows the same information as a directed graph of the same sub-population.

In a recent study Bell, Belli-McQueen and Haider (Bell, Belli-McQueen & Haider, 2007) argue that forgetting relationships can depend on network size as well as on the distance be-

tween a specific event and the recall of those memories. It is easier to remember the contact with only two former colleagues in the last year than contacts to a much larger number. Furthermore, the occasion or event where former colleagues met also plays an important role. In the underlying feasibility study, the respondents had to report their individual degree or personal network size in the target population during the last year (see Section 2). Although the requested time frame is quite long compared to other RDS studies, it seems to be appropriate for that special population. In the questionnaire distributed together with the network analysis, respondents reported several annual events they still attend or organize together with former colleagues (e.g., summer barbecues, Christmas parties or biker meetings). Those convivial gatherings help strengthen former corporate network structures and thus ensure that the network remains dense enough to use the RDS methodology. Furthermore, those informal meetings lead to the creation of some kind of hierarchical shadow, where information about different facets of the termination or re-placement process are spread between survivors and victims of the relocation (e.g., the fairness of layoff procedures or the adequacy and results of outplacement services). Future research efforts are needed to evaluate the influence of masking and forgetting to report relationships on RDS. Wejnert concludes that inferences on relative network differences remain valid if the *"recall error in self-report network data is independent of most key respondent attributes and demographics"* (Wejnert, 2010, p. 122).

Figure 7: Memory Gaps in Degree Reporting and/or Masking of Relationships



Source: Own work analyzed with Visone (http://www.visone.info).

6. Concluding Remarks

This paper shows that Respondent-Driven-Sampling can be an appropriate research method to study hidden populations beyond traditional AIDS prevention or surveillance projects. De-

spite the limitations presented above, RDS can be used to get a fast and cost-efficient access to the hidden target population of people affected from relocations. A great advantage of RDS is that respondents do not need to provide names or contact information of potential participants as it is the case in traditional snowball sampling. Consequently, confidentiality for the respondents and their recruits is ensured and their participation in the study becomes entirely voluntary. According to Magnani et al. (2005, p. 70) the method produces also more reliable data with a greater external validity, since it is not limited to visible members of the target population.

Despite numerous advantages of RDS presented in this paper, the method has also various limitations. Recent research indicates a high sensitivity to the assumptions presented in Section 2 (Gile & Handcock, 2010; Mouw & Verdery, in press; Verdery & Mouw, in press). Other issues are the high sensitivity to missing data, to clustering and to social homophily (Verdery & Mouw, in press; Wejnert, 2010) . Thus, it is often recommended to use a 'design effect' of at least 2.0, which would mean to recruit twice as many individuals as would be needed in a simple random sampling approach (see Salganik, 2006, p. 106). In a recent simulation study, Goel and Salganik (2010) used 85 real populations, where the network structures and demographic compositions are known, to evaluate the performance of RDS. They found a much higher 'design effect' with an overall median 'design effect' across all 85 simulated studies of 5.9 (Goel and Salganik, 2010, p. 6744). To use RDS, it is therefore suggested to find as much information as possible about the true demographic compositions (e.g., the ratio of blue- and white-collar victims) inside the target population to verify the results found with the social network recruiting approach of RDS.

Besides the methodological findings presented in this paper, the results also provide some new insights and starting points that could contribute to a deeper understanding in the analysis of transaction costs economics. Section 5 illustrated the importance of hubs or sociometric stars that function as bridges between victims and survivors of relocations. Those highly connected hubs play an important role for the spread of information or rumors between those two sub-groups, leading to the formation of a hierarchical shadow, where former organizational structures and information flows remain stable over time. This has some important implications for companies trying to reduce their costs through relocations.

Husted and Folger (2004) identified an important set of transaction costs in their work that they termed 'fairness response transaction costs'. Given the dynamic nature of the perception of fairness the authors argue further *"that organizational justice* [literature] *suggests that governance mechanisms selected ex ante may fall out of alignment during the course of transaction"* (Husted & Folger, 2004, p. 726). In terms of relocation decision, this means that the offer of outplacement services or the delegation to transfer or transitional companies during the termination process may initially (ex ante) be viewed as fair. However, Husted and Folger state that *"unforeseen contingencies or disturbances may arise that require further bargaining, monitoring, and enforcement costs in order to reassure the parties that the exchange continues to be fair"* (ibid., p. 725). By failing to take into account those ex post transaction costs, the selected governance mechanism (e.g., the delegation to a transitional

company - often restricted to one year in Germany) may increase transaction costs due to responses to perceived injustice (e.g., remaining unemployment after one year in a transfer company).

Due to the hierarchical shadow, the perceived injustice of relocation victims will also affect informal aspects (e.g., psychological contracts) of the transaction relationship survivors of the relocation have with their employer. Companies that relocate their workforce should therefore monitor and actively manage their laid-off employees even years after the initial event. Reassuring that those victims maintain a kind of post-relocation affective commitment will have important implications to questions like the willingness of re-employment, organizational endorsement and the prevention of negative marketing of terminated employees as well as on the attitudes and behaviors of the remaining workforce. Nevertheless, these aspects have been insufficiently addressed in transaction cost analyses so far.

7. References

Baudry, B. & Chassagnon, V. (2010). The close relation between organization theory and Oliver Williamson's transaction cost economics: A theory of the firm perspective. Journal of Institutional Economics, 6(4), 477-503.

Bell, D. C., Belli-McQueen, B., & Haider, A. (2007). Partner naming and forgetting: Recall of network members. Social Networks, 29(2), 279-299.

Berg, A. (2011, April 13). Social networks in Germany. [Soziale Netzwerke in Deutschland]. Press Conference of the German Federation of the Information Industry, Telecommunications and New Media (BITKOM), Berlin, Retrieved May 2, 2011, from http://www.bitkom.org/files/documents/PK_Praesentation_Social_media.pdf.

Coryn, C. L., Gugiu, P. C., Davidson, E. J., & Schröter, D. C. (2007). Needs assessment in hidden populations using respondent-driven sampling. Evaluation Journal of Australasia, 7(2), 3-11.

Destatis (2008). Relocating economic activity - results of a pilot survey [Verlagerung wirtschaftlicher Aktivitäten - Ergebnisse der Piloterhebung]. Wiesbaden: German Federal Statistical Office.

Gabor, M. R. (2007). Types of non-probabilistic sampling used in marketing research. Snowball sampling. Management & Marketing, (3), 80-90.

Gandolfi, F. (2006). Corporate downsizing demystified: A scholarly analysis of a business phenomenon. ICFAI Books.

German Federal Statistical Office and statistical Offices of the Länder (2010). Kreiszahlen: Ausgewählte Regionaldaten für Deutschland (2009 ed.). Hannover: Landesbetrieb für Statistik und Kommunikationstechnologie Niedersachsen (LSKN).

Gile, K. J. & Handcock, M. S. (2010). Respondent-Driven sampling: An assessment of current methodology. Sociological Methodology, 40(1), 285-327.

Goel, S. & Salganik, M. J. (2009). Respondent-Driven sampling as markov chain monte carlo. Statistics in Medicine, 28(17), 2202-2229.

Goel, S. & Salganik, M. J. (2010). Assessing respondent-driven sampling. Proceedings of the National Academy of Sciences (PNAS), 107(15), 6743-6747.

Goodman, L. A. (1961). Snowball sampling. Annals of Mathematical Statistics, 32(1), 148-170.

Hagan, J. (2001). Cause and country: The politics of ambivalence and the american vietnam war resistance in canada. Social Problems, 48(2), 168-184.

Heckathorn, D. D. (1997). Respondent-Driven sampling: A new approach to the study of hidden populations. Social Problems, 44(2), 174-199.

Heckathorn, D. D. (2002a). Development of a theory of collective action: From the emergence of norms to AIDS prevention and the analysis of social structure. In J. Berger & M. Zelditch (Eds.), New directions in contemporary sociological theory (J. Berger & M. Zelditch, Eds.). (pp. 79-108). Lanham, Md [u.a.]: Rowman & Littlefield Publ.

Heckathorn, D. D. (2002b). Respondent-Driven sampling II: Deriving valid population estimates from chain-referral samples of hidden populations. Social Problems, 49(1), 11-34.

Heckathorn, D. D. (2007). Extensions of respondent-driven sampling: Analyzing continuous variables and controlling for differential recruitment. Sociological Methodology, 37(1), 151-207.

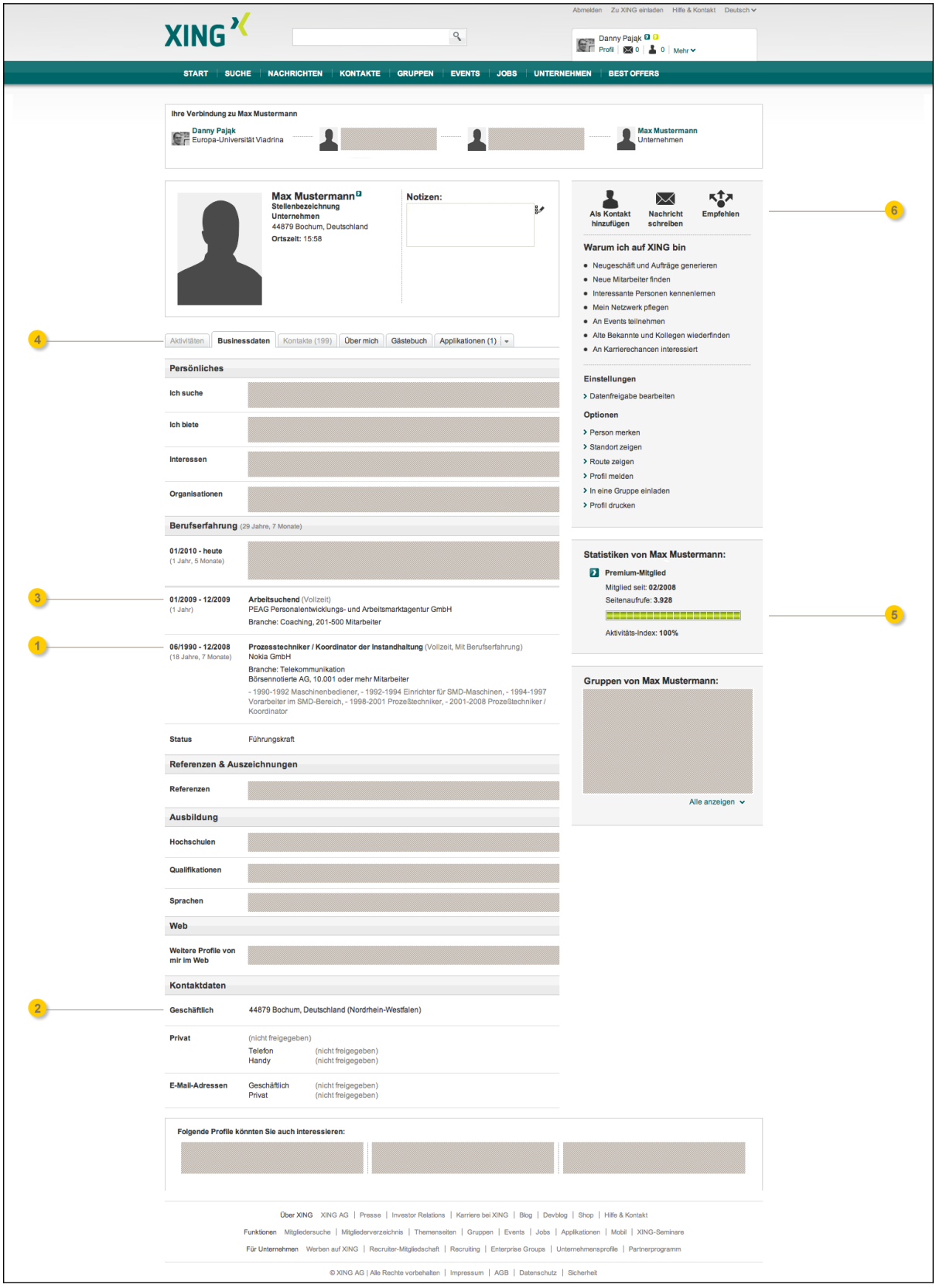
Heckathorn, D. D. & Jeffri, J. (2001). Finding the beat: Using respondent-driven sampling to study jazz musicians. Poetics, 28(4), 307-329.

Heckathorn, D. D. & Jeffri, J. (2003). Social networks of jazz musicians. Changing the Beat:

A Study of the Worklife of Jazz Musicians, Volume III: Respondent-Driven Sampling: Survey Results by the Research Center for Arts and Culture, National Endowment for the Arts Research Division Report, (43), 48-61.
Heckathorn, D. D. & Jeffri, J. (2005, June 22). Assessing the feasibility of respondent-driven sampling: Aging artists in new york city. Working paper [Web page]. Retrieved November 17, 2008, from http://www.tc.columbia.edu/centers/rcac/pdf/FeasRep_12.pdf
Heckathorn, D. D., Semaan, S., Broadhead, R. S., & Hughes, J. J. (2002). Extensions of respondent-driven sampling: A new approach to the study of injection drug users aged 18--25. AIDS and Behavior, 6(1), 55-67.
Husted, B. W. & Folger, R. (2004). Fairness and transaction costs: The contribution of organizational justice theory to an integrative model of economic organization. Organization Science, 15(6), 719-729.
Jackson, M. O. (2006). The economics of social networks. In R. Blundell, W. Newey, & T. Persson (Eds.), Advances in economics & econometrics, theory and applications: Ninth world congress of the econometric society (R. Blundell, W. Newey & T. Persson, Eds.). (pp. 1-56). Cambridge University Press.
Johnston, L. G. (2008). Behavioural surveillance - introduction to respondent driven sampling: Paticipant manual [Web page]. United States Department of Health and Human Services, Centers for Disease Control and Prevention (HHS-CDC). Retrieved August 27, 2009, from http://globalhealthsciences.ucsf.edu/PPHG/assets/docs/respondent-driven-sampling-2008.pdf
Johnston, L. G. & Sabin, K. (2010). Sampling hard-to-reach populations with respondent driven sampling. Methodological Innovations Online, 5(2), 38-48.
Johnston, L. G., Khanam, R., Reza, M., Khan, S. I., Banu, S., Alam, M. S., et al. (2008). The effectiveness of respondent driven sampling for recruiting males who have sex with males in dhaka, bangladesh. AIDS and Behavior, 12(2), 294-304.
Kinkel, S., Dachs, B., & Ebersberger, B. (2007). Production relocations and backshorings from a european comparative perspective. [Produktionsverlagerungen und Rückverlagerungen im europäischen Vergleich] Industrie-Management, 23(1), 47-51.
Magnani, R., Sabin, K., Saidel, T., & Heckathorn, D. (2005). Review of sampling hard-to-reach and hidden populations for HIV surveillance. AIDS (London, England), 19(2), 67-72.
Milgram, S. (1963). Behavioral study of obedience. The Journal of Abnormal and Social Psychology, 67(4), 371-378.
Mouw, T. & Verdery, A. M. (In press). Network sampling with memory. Currently Revise and Resubmit at Sociological Methodology. Retrieved May 17, 2011, from http://dupri.duke.edu/pdfs/mouw_2010_Network_sampling.pdf
Newman, M. E. J. (2005). Power laws, pareto distributions and zipf's law. Contemporary Physics, 46, 323-351.
Ramirez-Valles, J., Heckathorn, D. D., Vázquez, R., Diaz, R. M., & Campbell, R. T. (2005). From networks to populations: The development and application of respondent-driven sampling among idus and latino gay men. AIDS and Behavior, 9(4), 387-402.
Ravasz, E. & Barabási, A. -L. (2003). Hierarchical organization in complex networks. Physical Review E, 67(026112).
Réka, A., Jeong, H., & Barabási, A. -L. (2000). Error and attack tolerance of complex networks. Nature, 406(27), 378-382.
Salganik, M. J. & Heckathorn, D. D. (2004). Sampling and estimation in hidden populations using respondent-driven sampling. Sociological Methodology, 34(1), 193-240.
Schnettler, S. (2009a). A small world on feet of clay? A comparison of empirical small-world

studies against best-practice criteria. Social Networks, 31(3), 179-189.
Schnettler, S. (2009b). A structured overview of 50 years of small-world research. Social Networks, 31(3), 165-178.
Tiffany, J. S. (2006). Respondent-Driven sampling in participatory research contexts: Participant-Driven recruitment. Journal of Urban Health, 83(7), 113-124.
Verdery, A. M. & Mouw, T. (In press). Estimated sampling variance in respondent driven sampling: Do you know if you have an imprecise sample? Currently Under Review at Sociological Methodology. Retrieved May 17, 2011, from http://www.tedmouw.info/verdery%20and%20mouw%202011%20estimated%20sampling%20variance%20in%20respondent%20driven%20sampling.pdf
Volz, E. & Heckathorn, D. D. (2008). Probability based estimation theory for respondent driven sampling. Journal of Official Statistics, 24(1), 79-97.
Volz, E., Wejnert, C., Degani, I., & Heckathorn, D. D. (2007). Respondent-Driven sampling analysis tool (RDSAT) version 5.6. Ithaca, NY: Cornell University.
Wang, J., Carlson, R. G., Falck, R. S., Siegal, H. A., Rahman, A., & Li, L. (2005). Respondent-Driven sampling to recruit MDMA users: A methodological assessment. Drug and Alcohol Dependence, 78(2), 147-157.
Wang, J., Falck, R. S., Li, L., Rahman, A., & Carlson, R. G. (2007). Respondent-Driven sampling in the recruitment of illicit stimulant drug users in a rural setting: Findings and technical issues. Addictive Behaviors, 32(5), 924-937.
Wejnert, C. (2009). An empirical test of respondent-driven sampling: Point estimates, variance, degree measures, and out-of-equilibrium data. Sociological Methodology, 39(1), 73-116.
Wejnert, C. (2010). Social network analysis with respondent-driven sampling data: A study of racial integration on campus. Social Networks, 32(2), 112-124.
Wejnert, C. & Heckathorn, D. D. (2008). Web-Based network sampling: Efficiency and efficacy of respondent-driven sampling for online research. Sociological Methods & Research, 37(1), 105-134.
Williamson, O. E. (1976). The economics of internal organization: Exit and voice in relation to markets and hierarchies. The American Economic Review, 66(2), 369-377.

Appendix 1: Xing Sample Profile



Source: Xing (http://www.xing.com) - Modified and anonymized profile screenshot, retrieved May 9, 2011.