



Contents

Number 6, Issue 2, May 2014

Mohammad Morad, Shahabul Haque, Jahangir Alam

Contextualizing Formation of Diaspora of Bangladeshi Immigrants in the UK..... 103

Johannes Zück

Social change and economic transformations. Can variations in social capital across time affect the structure of an economy? 130

Marija Paladin

Informational value of data about data in surveys (Example of two web surveys)..... 154

RSC, Number 6, Issue 2, May 2014, pp. 103-128.

Contextualizing Formation of Diaspora of Bangladeshi Immigrants in the UK

Mohammad Morad, MSS, MA (Corresponding author)

Assistant Professor, Department of Sociology, Shahjalal University of Science & Technology, Sylhet, Bangladesh*

Md. Shahabul Haque, MSS, MA

Assistant Professor, Department of Political Studies, Shahjalal University of Science & Technology, Sylhet, Bangladesh

Md. Jahangir Alam, PhD

Associate Professor, Department of Sociology, University of Dhaka, Bangladesh

*Corresponding author's e-mail: moradsust@yahoo.com

Abstract: *In this study we contextualize the diaspora formation of Bangladeshi immigrants in the UK by analyzing three broad types of diaspora characteristics- dispersion; connection with the homelands; and maintaining a distinctive identity in the host society. Based on secondary data, this study revealed that Bangladeshi diaspora members in the UK are mainly dispersed from their country of origin for economic reasons. As many other diaspora group, Bangladeshi immigrants in the UK maintain several linkages- economical, social and cultural, and political- with their home country Bangladesh. In this host society, they present their distinctive life by maintaining close ties with ethnic Bangla cultural objects. They are also present their vibrant and distinctive Banglanness by doing several activities collectively through associations. By doing so, Bangladeshi immigrants in the UK work like a diaspora as the three broad elements of diaspora were clearly evident in our findings and discussion.*

Keywords: Diaspora, Bangladeshi Immigrants, Dispersion, Homeland Linkage, Distinctive Identity Maintenance

Introduction

In the era of globalization, diasporas have become an important subject of study among researchers, academics and scholars. The past decade has seen a variety of literature on this issue which introduced many academic, political and policy debates that spread across the discipline. In present times, it has become an academic area of research not only in Migration Study, but also in many disciplines including Sociology, Anthropology, History, Geography, Cultural Studies, Political Science and Literature. Several theorists such as Safran (1991), Cohen (1997), Vertovec (1999), and Brubaker (2005) have proposed criteria to define a diaspora. Their research has shown that the concept of diaspora has attained a broad semantic field and its meaning and uses have been proliferated in verity of directions to get the dispersion of this term. While the classical meaning of diaspora involves an involuntary migration, this term has been used for both voluntary and involuntary migration in the current discussion. Thus, in contemporary times, the meanings and uses of diaspora have been proliferated, where the term often implies the ongoing relationship between immigrants' homeland and their host countries (Verhulst, 1999:30). The current members of diaspora include overseas, ethnic, exile, minority, refugee, expatriated people, migrants and so on (Inbom, 2003:10).

Scholars have classified immigrants as members of a diaspora based on their ongoing relationship with their country of origin, in many of their contemporary studies the term 'diaspora' has often been used interchangeably with the term 'transnationalism' or 'transnational communities' who maintain social, emotional and political network that cross the borders of nation-states (Vertovec, 1999). This can also be discerned from Tölölyan's (1991: 5) statement indicating that contemporary diasporas are 'the exemplary communities of the transnational moment' (cited in Faist, 2010: 16). Thus, diaspora members maintain close linkages with their homeland and many of them have important and durable relationships, including socioeconomic,

cultural and political relationships, that flourish in two or more societies at once (Castles and Miller, 2009:3).

This paper attempts to make in the understanding of the diaspora formation of Bangladeshi immigrants in the UK by analyzing three broad categories of diaspora characteristics: dispersion; connection with the homelands; and maintaining a distinctive identity in the host society. Indeed, international migration has become a well-known phenomenon in Bangladesh; a large number of Bangladeshis are migrating outside of Bangladesh every year as long-term migrants. Among the total number Bangladeshis living abroad, according to Siddiqui (2004), there are almost 1.2 million that reside permanently as citizens, or with other valid documents in Western Industrialized countries of Europe, North America and also in Australia. In Europe the highest number of Bangladeshi migrants (500000) is found in the UK, followed by Italy which has 70,000 Bangladeshi, and Greece with 11,000 Bangladeshi (ibid), who maintaining several relations and contribute to the development of their homeland in several ways. In this paper, we first contextualize the term diaspora by focusing on its meaning and developing a theoretical framework. We then go on to outline the method of the study. Finally, diaspora formation of Bangladeshi immigrants in the UK is discussed on the basis of the outlined theoretical framework.

Diaspora and Its Formation: Theoretical Background

The Notion of Diaspora: An Introduction

Historically, there are three dimensions of diasporas: original, classical and contemporary (Inbom, 2003: 9). Originally, the term diaspora is derived from the Greek word '*diaspeiro*' which was used as early as the fifth century B.C. by Greek legends including, Sophocles, Herodotus, and Thucydides (Dufoix, 2007:4). Here, *Speiro* means 'to show' and *Dia* means 'over', hence, *diaspeiro* was used in ancient Greece to mean migration and colonization (Cohen, 1997: ix).

On the other hand, the classical discussions of diaspora were mostly rooted with homeland that was basically concerned about paradigmatic cases including Jews and some other 'classical' diasporas (Brubaker, 2005:1). Thus, in the classical notion, diaspora refers to the forced dispersion that is basically concerned about the traumatic history of dispersal, myths and memories of the homeland. For instance, Armenian, African and Palestinian along with Jews and some others who forcibly dispersed, such as the Irish¹, were referred to as a diaspora due to their collective trauma and fate to live in exile (Cohen, 1997).

In contemporary times, the specific definition of diaspora has been proliferated by academics and social scientists through their work. Research findings shows that the contemporary concept of a diaspora is a way of understanding migration, immigrants, any kind of dispersal, their identity maintenance, integrations, and transnational linkages, etc. In this vein, Richard Marienstras argues that diaspora concepts are increasingly used to explain any group or population that in one way or another has a history of migration (cited in Wahlbeck, 1998:10). In that way, numerous contemporary studies conceptualize different immigrant groups as diasporas. For instance, according to Anderson (1998), migrant groups who are defined as 'long distance nationalist' belong to diasporas because of their continuous involvement in home politics and nationalist movements, examples of these are the Kashmiri, Palestinians, Tamil and others (cited in Brubaker, 2005:2). William Safran's also argue that Cubans and Mexicans in the USA, the overseas Chinese, Poles, Palestinians and blacks in North America and the Caribbean, Turks in Germany, Indians and Armenians in different countries, Pakistanis in Britain, Maghrebis in France, and some others, are various diaspora groups around the world (Safran, 1991:83). Drawing on the same logic, Sheffer (2003) identifies a number of labor migrants as diaspora because of their emotional and social ties with their homeland. For instance, Bangladeshi, Pilipino, Indian, Greek,

¹ According to Cohen (1997:27) "The migration of the Irish over the period 1845 to 1852, following the famine, can be regarded as an analogous trauma".

Haitian, Italian, Korean, Mexican, Turkish, Polish, Salvadorian, Pakistani, Vietnamese, and many other labor migrants could be argued to be a diaspora (cited in Brubaker, 2005:2)

However, Cohen (1997) offered a more nuanced typology [Table 1] of diasporas on the basis of conditions of migration, using categories such as victim or refugee diaspora, labor or service diasporas, trade or business diaspora, imperial or colonial diaspora, and cultural or hybrid diaspora. According to Cohen, each of the diaspora groups have their own condition of migration, for instance, victim diasporas are created through forced dispersion from their homeland as a result of political unrest or persecution, whereas labor diasporas are formed by people leaving the homeland in search of work abroad. Cohen highlights that one diaspora group may belong to more than one category depending on their origin, function and characteristics. For example, Table I shows that during the colonial period Indians were indenture laborers but in the modern global world they became a trade/business and professional diaspora.

Table 1: Different Types of Diaspora

Types of Diaspora	Examples
Victim/Refugee	Jews, African, Armenians, Irish, and Palestinian
Imperial/Colonial	Ancient Greek, British, Russian, Spanish, Portuguese, Dutch
Labour/Service	Indentured Indian, Chinese and Japanese, Sikhs, Turks, Italians
Trade/business/professional	Venetians, Lebanese, Chinese, today's Indians and Japanese
Cultural/hybrid/postmodern	Caribbean and today's Chinese and Indian

Source: Cohen (1997:178)

Therefore, building on contemporary meanings from the above discussion, this study has chosen to define the term diaspora as a group of immigrants who maintain various relationships with their homeland.

Formation of Diaspora: Theoretical Framework

Several theorists have proposed criteria for a diaspora's characteristics. In order to present the theoretical framework of this paper here we present the prominent diaspora paradigm that is discussed by Safran (1991) and Cohen (1997). In addition, two contemporary discussion of diaspora formation, those of Vertovec (1999) and Brubaker (2005) are also addressed.

First, Safran (1991: 83-84) propose some component of diaspora in his study- '*Diasporas in Modern Societies: Myths of Homelands and Return*'. Here diaspora have been defined as expatriate minority communities that contain following key components:

1) They, or their ancestors, have been dispersed from a specific original "center" to two or more "peripheral," or foreign, regions; 2) they retain a collective memory, vision, or myth about their original homeland—its physical location, history, and achievements; 3) they believe that they are not—and perhaps cannot be—fully accepted by their host society and therefore feel partly alienated and insulated from it; 4) they regard their ancestral homeland as their true, ideal home and as the place to which they or their descendants would (or should) eventually return—when conditions are appropriate; 5) they believe that they should, collectively, be committed to the maintenance or restoration of their original homeland and to its safety and prosperity; and 6) they continue to relate, personally or vicariously, to that homeland in one way or another, and their ethnocommunal consciousness and solidarity are importantly defined by the existence of such a relationship. (Safran, 1991:83-84)

Second, another list of key features of diaspora is found from the discussion of Cohen; he uses Safran's criteria, but supplements some points. He merges

criteria number four and five of Safran, i.e., *'their ancestral homeland as their true, ideal home'* with *'committed to the maintenance or restoration'* into one and adds a line *'even to its creation'*. On the other hand, he adds four more criteria to that of Safran: *groups that disperse for colonial or voluntarist reasons; a diasporic identity; mobilize a collective identity; in solidarity with co-ethnic members in other countries* (Cohen, 2008:6-7). As a result, Cohen (2008) has listed the following nine components for defining diaspora:

1. Dispersal from an original homeland, often traumatically to two or foreign regions; 2. Alternatively or additionally, the expansion from a homeland in search of work, pursuit of trade or to further colonial ambitions; 3. a collective memory and myth about the homeland including its location, history, suffering and achievements; 4. an idealization of the real or imagined ancestral home and a collective commitment to its maintenance, restoration, safety and prosperity, even to its creation; 5. the frequent development of a return movement to the homeland that gains collective approbation even if many in the group are satisfied with only a vicarious relationship or intermittent visits to the homeland; 6. a strong ethnic group consciousness sustained over a long time and based on a sense of distinctiveness, a common history, the transmission of a common cultural and religious heritage and the belief in a common fate; 7. a troubled relationship with host societies, suggesting a lack of acceptance or the possibility that another calamity might befall the group; 8. a sense of empathy and co-responsibility with co-ethnic members in other countries of settlement even where home has become more vestigial; and 9. the possibility of a distinctive creative, enriching life in host countries with a tolerance for pluralism. (Cohen, 2008:17).

Third, analyzing Sarfen and Cohen and, some other proponents, Brubaker (2005:5) has identified the three following criteria for diaspora which he mentioned as core elements for diaspora formation:

1. Dispersion: it refers to any kind of dispersion in space that crosses state borders. According to him it is widely accepted but not a universal criterion (ibid).

2. Homeland Orientation: the second criterion is homeland orientation. Here he includes four of the six criteria of Safran related to homeland orientation such as, (i) maintain a memory, vision, or myth about the homeland; (ii) ancestral home as a place of eventual return to homeland; (iii) committed to the maintenance and restoration of this homeland; and (iv) a continuing relationship with the homeland (ibid).

3. Boundary-Maintenance: it is involved with the preservation of distinctive identity where one can be part of a diaspora with the following distinctive and relatively dense social relationship as a transnational community through their links that cross state boundaries (ibid, 6).

Fourth, considering the transnationalism of diaspora, Vertovec (1999) conceptualizes the term diaspora *as a social form, as a type of consciousness, and as a mode of cultural production.*

1) Diaspora as a social form, he addresses the social, economic and political networks of the transnational communities. In this notion, he includes three social categories. The first category is 'a specific kind of social relationship' where the diaspora is seen as consequences of voluntary or forced migration; consciously maintaining collective identity; institutionalizing networks of exchange and communication; maintaining a variety of explicit and implicit ties with their homelands; developing solidarity with co-ethnic members; inability or unwillingness to be fully accepted by 'a host society' with feelings of alienation, exclusion or others differences (Vertovec, 1999). The second category is called '*political orientation*' where individual immigrants are significant actors who through their collective organization work as a pressure group in the domestic politics of their host country for the favor of their country of origins (ibid,4). The third category is 'economic strategies of transnational groups' which he mentions as an important source and force of international finance (ibid).

2) Diaspora as a type of consciousness, he refers to a particular type of awareness that is based on a variety of experiences which generated among the contemporary transnational communities (ibid).

3) Diaspora as a mode of cultural production, he refers to production and reproduction of transnational social and cultural phenomena which constructed the styles and identities of diaspora communities (ibid).

In this vein, it seems clear that there are no universal characteristics of a diaspora. Thus, the elements considered to define an immigrant community as diaspora mainly depends on its dispersal condition and homeland connections. However, it seems from the above discussion that there is an understanding among the authors about three broad types of elements of diaspora. These include (1) dispersion: which can be voluntary and non voluntary; (2) connection with the homelands; (3) distinctive identity maintenance in the host society. Therefore, this study will use these three broad aspects of diaspora characteristics: dispersion; homeland connection; and maintaining a distinctive identity, in order to better understand the formation of the Bangladeshi diaspora in the UK.

Methodology

The research methodology for this article involved rigorous analysis, and the examination and evaluation of literature relevant to the diaspora formation of Bangladeshi immigrants in the UK. Therefore, this paper on the one hand analyses the theoretical discussion about the concept of diaspora and its formation based on several authors' discussion in different books, journals and academic papers. On the other hand, diaspora formation of Bangladeshi immigrants is conceptualize through the analysis of secondary sources that include several articles, research papers, reports of both governmental and non-governmental organizations about Bangladeshi migration to the UK and their homeland relations and their identity maintenance.

Result and Discussion: Formation of Bangladeshi Diaspora in the UK

Desperation

We have already showed in the theoretical discussion that the older notion of a diaspora was mainly concerned on the forced dispersal of a people, such as the Jews, the modern notion of this term refers to any kind of dispersal. For instance, Cohen (1997) has proposed trade diasporas and labor diasporas, whose dispersion happened mainly for economic reasons. In this study, to investigate the nature of dispersion of the Bangladeshi migrants in the UK, the following section focuses on the history of Bangladeshi migration to the UK.

It is claimed that Bangladeshi migration to the UK is connected to Bangladeshi's history of British Colonialism (Siddiqui, 2004). As Adams (1994) points out, many Bangladeshis decided to migrate to the UK in search of work as Bangladesh was a part of the British Empire (cited in Hussain and Mirza, 2012: 74). Indeed, Alam stated (1988, cited by Siddiqui, 2004: 17) that part of landless people from the *Sylhet* region (the north east area of modern day Bangladesh) found jobs as dockyard workers, cooks, cook-mates or cleaners over the late nineteenth and early twentieth century in British merchant navy ships that carried goods from Kolkata in India to all other parts of the world. This group of people did not have much seafaring experiences. For that reason, according to experts, they left ships when presented with the opportunity (ibid) and found themselves in a number of countries, such as the USA and the UK, where they introduced many small settlements (Sikder, 2008:258). From this group of people of the *Sylhet* region, Bangladeshis who jumped ship in the UK, mostly settled in London, Liverpool and Bristol from the 1850s onwards (Change Institute, 2009) and sought work as peddlers, or in hotels and restaurant (Gardner, 2006). This people (i.e.: the *Sylheti*) are identified as the main pioneers of the British Bangladeshi diaspora. Still, the majority of Bangladeshis in the UK come from the region of Sylhet (Garbin, 2009:2).

The second wave of Bangladeshi migration to UK, started in 1950s and 1960s (Siddiqui, 2004; Gardner, 2006, and Change Institute, 2009) since UK conceived a new policy to encourage labor migration from its former colony due to its labor shortages. UK's Bangladeshi diaspora members took advantage of this opportunity to help their kin and kith to migrate to the UK by providing them with credit, arranging documents and thus gradually spreading their network (Gardner, 1993). As a result, a large number of Bangladeshis, mostly from *Sylhet*, arrived in the UK and most of them started a life as laborers in the heavy industries of Birmingham and Oldham, a few were settled in London and continued their trade as tailors (Siddiqui, 2004). Researchers explained this immigration as a long history of male immigration because it was not until the 1970s that their family members (i.e.: wife and children) joined them in the UK. Family reunifications began during the late 1960s and peaked in the 1970s (Change Institute, 2009:25). Beside economic migrants, according to Siddiqui (2004), a small number of non economic migrants, highly educated people from the upper and middle class of Bangladesh, also migrated to the UK before the Second World War to pursue higher education. However, compare to the economic migrants, this number was relatively small (ibid).

Thus, from the above history of Bangladeshi migration to the UK, it can be said that Bangladeshi immigrants in the UK are mainly dispersed from their country of origin for economic reasons (i.e. to earn a living). Thus, the nature of their dispersion is similar in nature with Cohen's labor diaspora - "*a diaspora can be generated from by emigration in search of work*" (Cohen, 2008: 61).

Connection with the Homeland

Several authors (e.g. Safran, 1991; Cohen, 2008; Brubaker, 2005; and Vertovec, 1999) have identified 'homeland connection' as an indispensable criterion for diaspora formation. Many studies have shown that as a dispersed community, Bangladeshi migrants in the UK maintain various ties with their family, relatives, and friends who live in Bangladesh. They have established a strong social, economic, cultural, and political relation with their homeland. As it is highlighted in one of the recent studies:

“The community [Bangladeshi in the UK] maintains strong ties with Bangladesh through travel, remittances, trade and commerce, cultural ties and politics. These are stronger with the older generation and whilst many young British Bangladeshis still value Bangladesh as the country of their roots and heritage; few are willing to invest, send money regularly, or stay in the country for a long term” (Change Institute, 2009:6).

The following section present and discuss the findings related to the homeland linkages of Bangladeshi immigrants in the UK by dividing it into economical, social and cultural, and political linkages.

Economic Linkages: Garbin (2005) has argued that in material terms, the ties of Bangladeshis in the UK with their homeland are expressed by their sending of remittances to their families and relatives who remain in their country of origin. A recent study on the Bangladeshi diaspora in the UK and the US shows that 84 percent of the respondents from the UK were sending remittance to Bangladesh (Siddiqui, 2004: 47). Table 2 presents the share of remittances received from the UK among the total remittance to Bangladesh. The Table shows that the contribution of remittances from the UK among the total remittance received in Bangladesh was 8% (US\$ 789.65 million) in the financial year, 2008-09 and 7.53% (827.51 million) in financial year of 2009-10 and 7.63% (889.60 million) in 2010-11, 7.69% (987.46 million) in 2011-12, and 6.86% (991.59 million) in 2012-13. Among these financial years, according to the Central Bank of Bangladesh, the UK was the fifth largest sources of remittance next to Saudi Arabia, UAE, USA, and Kuwait since 2011-12. In the year 2012-13, it became the sixth largest remittance sending country next to Saudi Arabia, UAE, USA, and Kuwait and Malaysia.

Table 2: Share of remittances from the UK in terms of total remittances received in Bangladesh

Remittance Received (USD in millions)	Fiscal year				
	2008-09	2009-10	2010-11	2011-12	2012-13
Total remittance to Bangladesh	9689.26	10987.40	11650.33	12843.44	14461.13
Remittance from the UK	789.65	827.51	889.60	987.46	991.59
Share from the UK (%)	8%	7.53%	7.63	7.69	6.86

Source: Central Bank of Bangladesh, 2013

Studies show that these remittances are mainly used for the development of family dynamics such as for maintaining own or extended families expenditures, purchasing/constructing land and houses, and for increasing family income by developing business (Siddiqui, 2004). Occasionally remittances is also sent for the support to their non-migrant relatives and friends in Bangladesh that help them either survive or migrate (Siddiqui, 2004; Change Institute, 2009). Moreover, remittances are also send for the purpose of social welfare and religious activities in their natal villages that are used for charity, support during natural disasters, and for providing *Zakat*² (ibid).

² Zakat is the third Pillar of Islam, which is regarded as a type of worship and self-purification by practice of a mandatory offering of a set proportion of one's accumulated wealth to charity.

Besides, Eade and Garbin (2003:9) have observed that UK based remittances have a great impact on the local landscape of many *Sylheti* villages, where most of the Bangladeshis in the UK originate. For instance, in this region many new roads, modern houses, religious institutions like mosques, madrassahs, and educational institution like schools have been constructed with the remittances from the UK that changed the outlook of the rural landscape. This has made this area an exemption in terms of infrastructure when compared to other rural areas in Bangladesh. Therefore, the above mentioned Bangladeshi immigrants' economic findings are in high concordance with the diaspora features presented by Cohen (2008) and Vertovec (1999) stating that immigrants have a 'sense of empathy and solidarity with co-ethnic members' that live in their natal villages in their homeland.

With regard to investments, research findings show that many people from the Bangladeshi community in the UK were investing in various sector including hotels, private property, and food import businesses which they considered permanent assets as these could be used as security measures upon return to their country of origin (Change institute, 2009). Thus, the motivation behind the investments is indicative of the diaspora characteristic of Safran (1991): they regard their ancestral homeland as their true, ideal home and as the place to which they, or their descendants, will eventually return when conditions are appropriate. Besides individually sending money, Bangladeshi immigrants in the UK also sent remittances collectively through several associations that mainly based on the district, village and town of their local area of origin (Siddiqui, 2004). These remittances are mainly used for community development and charitable purposes such as fund raising for local schools, providing scholarship for poor students, building mosques, infrastructure repairs, relief activities during natural disasters, and other reconstruction activities (ibid). Similarly to Safran (1991) and Cohen, (2008) this collective initiatives indicates that UK based Bangladeshi immigrants are 'collectively committed to the maintenance or restoration of their original homeland'.

Social and Cultural Linkages: according to the Change Institute (2009) kinship is extremely important in the case of Bangladeshi immigrants in the UK. Immigrants maintain communication with and regularly visit their family members, friends, and relative who remain in their home country. The reasons for their visits varies largely but can be broadly categorized in the following groups: for charitable purposes, often after political, economic, and environmental crises; in order to arrange marriages/wedding ceremonies; and to bury their dead (ibid). It has been argued that the increased availability and affordability of telecommunications made the communication stronger and denser between Bangladeshi immigrants in the UK and Bangladesh (Zeitlyn, 2011). For instance, during the 1980s, transnational *Sylheti* families in the UK mainly communicate with their natal villages through letters, which took weeks or months to reach their destination, as most households did not have a telephone. However, now, due to the availability of mobile phones Bangladeshis in London are involved in the day-to-day affairs and decisions of their families in Bangladesh. They can also feel the sense of companionship with other large joint family members in Bangladesh that they miss through their frequent communication with relatives in Bangladesh (ibid).

Behind the motivation of the above mentioned personal contacts, it has also been found that Bangladeshi migrants want to foster strong ties between their new generations who are growing up in the UK and their family members who still live in Bangladesh through regular telephone conversations and visit. For instance, Mand (2010) worked with British-born Bangladeshi children aged 9 to 10 and finds that most of the children whom she interviewed maintain transnational social relationship between Bangladesh and the UK, thus making them active members of transnational families. These immigrants' children often visit their parents' home, the Sylhet region in Bangladesh twice, or even three times a year, most times along with their families. There they often attend major festivals and functions such as weddings. In addition, weekly conversation over the telephone with their family members in Bangladesh were found to be the most common way in which children kept in touch with their grandparent and other family members. For these social relationships,

children have portrayed Bangladesh as their *desh* (home) and Britain is *bidesh* (a country away from home). Therefore, Just as Cohen (2008) argues, we can argue that Bangladeshi immigrants in the UK constructed their homeland as the ideal way of maintain social relationships. Besides, as Cohen (2008:17) stated diasporas have “*a sense of empathy and co-responsibility with co-ethnic members in other countries of settlement*”, Garbin (2005) also noted that the UK based Bangladeshis also maintain transnational social ties with other Bangladesh diaspora elsewhere. This is sustained through their religious rituals; circulation of goods and gifts; communication over phone, e-mail, and other social media; matrimonial links, i.e., selection of partners from Britain, America or other parts of the world instead of Bangladesh.

On the other hand, it seems that Bangladeshi migrants in the UK have cultural attachments with their homeland. They maintain their cultural linkage by means of several ethnic newspapers and TV channels in this host country, the UK. For example, they introduced various electronic and print ethnic media such as Bengali news papers including *Janamat, Natun Din, Shurma, Patriaka, Sylheter Dak, and Euro-Bangla*; English news papers such as *Dainik Bangladesh*; other forms of media such as radio and television channels including ‘*Bangla TV UK*’ have also been introduced (Siddiqui 2004). According to Siddiqui (2004) these media provides necessary information related to several immigration issues, and it provides an update of day to day politics, economic, and social aspects of Bangladesh. These media also tend to be involved in lobbying and networking with the British Governments on behalf of the Bangladeshi community (ibid). Apart from these cultural linkages, they also celebrate different Bangladeshi national, traditional social, cultural and religious festivals in this host society. Immigrants do this collectively through associations which provide Bangladeshis with opportunities to gather and to build up social linkage between their community’s people who live and work in this country (ibid).

Furthermore, immigrants also remain connected to other Bangla traditions. For instance, findings of several researches show that they have

strong linkages with their Bengali language, even with the local dialect, *Sylheti*. According to the Health Survey for England in 1999 (cited in Change Institute, 2009), 54 per cent of the UK Bangladeshi age between 16 and over used Bengali as their main language, followed by *Sylheti* (25 per cent) and English (20 per cent). Besides, 90 per cent of Bangladeshi Londoners use Bengali as the main medium of communication in their home (Change Institute, 2009), while they use English mainly outside their home (Lawson and Sachdev, 2004). It is also evident that Bangladeshi people in the UK have linkages with their ethnic food. Throughout London, because of the large demand among the Bangladesh community, Bangladeshi spices, vegetables, fish, fruits, sweets, snacks are readily available in Bangladeshi shops and supermarkets (Jennings et al, 2014). Immigrants even maintain the tradition of chewing *Paan*, a mixture of betel leaf and areca nut, as is common in Bangladesh (Hussain and Mirza, 2012). Bangladeshi migrants' linkage with their ethnic food is also evident in their habits of bringing back a variety of Bangla foods when they return to the UK after visiting Bangladesh. As it is mentioned by the Gardner (1993:11):

“The bags of migrants returning to Britain are often filled with chutney, pickled mangoes, and dried fish. Once in Britain these are distributed to the kin of village neighbors living nearby. Likewise Bangladeshi food is readily available in British cities. In London, fresh Sylheti fish is flown in daily. In season jackfruits (selling for twenty or thirty pounds each) can be bought in Brick Lane and Spitalfields markets. Most families consume rice from Bangladesh or India, along with betel nut, spices and a wide variety of Bengali vegetables”.

Therefore, just as Cohen's (2008) proposed criteria it seems clear that they are presenting their 'distinctive creative, enriching life' in their host country, the UK by preserving and promoting their ethnic culture.

Political Linkages: it has been argued that since the 1960s, when a sizeable Bangla community began to emerge to the UK, the issue of home country's politics became an important issue among this community (e.g.

Garbin, 2008; Alexander and et al, 2010). Garbin (2009) has argued that first generation migrants were actively involved in the resistance leading to the 1971 Bangladesh Liberation War (*Muktio Juddo*) against Pakistan. These UK based migrants supported their home country's freedom fighters by fund raising for them, by holding public protest, and with their lobbying and networking with the British government, institutions, and media (Garbin, 2009; Alexander and et al, 2010). These activities have been considered as '*long distance nationalism*' of Bangladeshi diaspora that promoted a collective and authentic Bengali identity among the British Bangladeshi community which transcended class, caste, and religious boundaries (ibid).

According to Garbin (2009), political activities of Bangladeshi immigrants in the UK are sustained through a set of networks and practices that connects groups and local communities, as well as encourages the movement and circulation of people between Bangladesh and the UK. It has been argued that nearly all Bangladeshi political parties have been represented in Britain since the mid-1980s (Garbin, 2009). Most of the members of these political parties are first generation migrants who settled in the UK in 1960s along with other Bangladeshis who came during 1980s and 1990s. With regard to their activities, it has been shown that Bangladeshi political parties in this host society perform social and political work, mainly campaign for their parties during election times in Bangladesh. In addition, these parties help first generation migrants to get elected and develop economic activities and to protect business interests in Bangladesh (ibid).

On the other hand, studies showed that second generation Bangladeshis in the UK are mostly involved with UK's politics. The majority of this generation supports the Labor Party, and some members even became actively involved in local politics as elected members (Change Institute, 2009). For example, *Rushanara Ali*, a young British Bangladeshi, was elected as a member of the British Parliament as a British Labor Party representative in 2010. In light of the study of Vertovec (1999), however, it has been shown that by participating in a host country's electoral politics, Bangladeshi immigrants

in the UK are lobbying in the favor of their homeland by influencing policies in favor of Bangladesh (Siddiqui 2004:13).

Distinctive Identity Maintenance in the Host Society

The third element which has been considered in the investigation of diaspora formation is the 'distinctive identity maintenance in the host society' that several authors argued is an essential element for a diaspora (e.g. Safran, 1991:83; Cohen, 1997:24; Brubacker, 2005:6, Vertovec (1999:279). In this case, some features that have been outlined in the discussion of the theoretical framework were also evident in the case of Bangladeshi immigrants living in the UK.

In this regard, as Vertovec (1999) has argued, the collective identity of a diaspora is sustained by their common origin and historical experiences, this aspect is evident in the activities of Bangladeshi immigrants in the UK and it is also evident that they maintain this community and traditions in order to maintain their distinctive identity. For instance, focusing on this issue, the Change Institute (2009:39) has stated the following:

"They [respondents of Bangladeshi immigrants in the UK] suggest that this identity is strongly linked to their cultural heritage, their nationality of origin, and affinities with the politics and political parties of Bangladesh".

Indeed, in the UK, the distinct secular nationalist *Bangla* heritage is expressed by the visible symbols of *Shaheed Minar*, a monument which commemorates the martyrs (*shaheed*) of the Bangladesh's Language Movement of 1952, and the sculpture of the *Shapla*, the national flower of Bangladesh (Garbin, 2005; Eade and Garbin, 2006). The *Shaheed Minar* was erected in Altab Ali Park, Whitechapel, Tower Hamlet and in Oldham, that was built in the Bangladeshi area of Westwood along with the sculpture of the *Shapla* (ibid). The researchers highlight that for the Bangladeshi community, "*this visibility was a crucial marker of collective identity*" (Eade and Garbin, 2006, 186). It is also

worth mentioning that the days and festivals Bangladeshi migrants in the UK celebrate are related to the national history of their homeland. For instance, they celebrate the date of language movement (21 February), Independence Day (26 March) and Victory day (16 December) of Bangladesh (Siddiqui 2004). Thus, their collective *Bangla* identities originate from their common historical background. These examples indicate that Bangladeshi migrants in the UK are 'maintaining a strong ethnic group conciseness and collective identity' that being a key feature of Cohen's (2008) definition of a diaspora.

With regard to identity maintenance, another diaspora feature '*distinctive creative, enriching life in the host countries*' proposed by Cohen (2008) is also evident in the activities of Bangladeshi immigrants. As we have shown in the previous section, Bangladeshi migrants in this host society present their distinctive life by maintaining close ties with ethnic *Bangla* cultural objects. As we have already demonstrated, they maintained strong linkages with their ethnic media, food, and language with and some other cultural objects. They also present their vibrant and distinctive *Banglaness* by doing several activities collectively with their ethnic associations. For instance, we have already mentioned that Bangladeshi associations observe different home festivals. In these occasions, they follow their traditional rituals, as is done in Bangladesh. For example, different types of homemade food and small handcrafts stalls are arranged, and several cultural programs are staged where different Bangladeshi artists performed traditional Bangle folk music, dance, drama and pop music (Siddiqui, 2004; Eade and Garbin, 2006). In addition, programs in the national days are also framed with the Bangladeshi national anthem, speeches, and recitation of Bengali poems. They also arrange several musical shows where popular artists from Bangladesh perform (Siddiqui, 2004; Eade and Garbin (2006)). It has been argued that through these activities aimed at preserving and promoting *Bangla* culture, they made Brick Lane a "Bangla town" where the 'International Curry Festival' and the celebration of *Baishaki Mela* (the Bengali New Year) play important roles (Eade and Garbin, 2006).

Therefore, these migrant activities- individual practices and associational activities- indicate that Bangladeshis are presenting their distinctive identity through their cultural practices in their host society, the UK, and at the same time this collective activities indicates their 'collective commitment to the maintenance of homeland' (Safran, 1991, Cohen, 2008).

Conclusion

This paper has attempted to contextualize the diaspora formation of Bangladeshi immigrants in the UK, with a focus on their nature of dispersion, connection with the homeland, and distinctive identity maintenance in the host society. First, with regard to the nature of the dispersion, the findings of this study indicate that Bangladeshi immigrants in the UK are similar in nature with Cohen's definition of a labor diaspora - a diaspora can be generated by emigrant in search of work (Cohen, 2008). Second, with regard to the homeland linkages, as Vertovec (1999:279) mentioned, diasporas maintain '*a variety of explicit and implicit ties with their homelands*', taken as a whole, the above discussion of the section homeland linkages has delineated the fact that, UK based Bangladeshi migrants have been playing crucial roles in order to maintain linkages with their country of origin. They have built up social and economic linkages with families, relatives, and friends that they left in their home country. Also they have maintained cultural linkages with their homeland by retaining their cultural roots, which is visible through their attachment with ethnic media, i.e., *Bangla* television channels and newspapers, with ethnic food in the daily meals, and activities in Bangladeshi associations, and participation in national and cultural festivals of Bangladesh. Besides, members of the first generation remain actively concerned with their home country's politics, while members of the second generation have involved themselves with local and national politics of their host country, the UK. It has been found out that nearly all the Bangladeshi political parties have been represented in the UK where most of the members of these political parties are first generation migrants who settled in the UK in 1960s, along with other Bangladeshis who came during 1980s and 1990s. In addition, by

participating in the host country's electoral politics, Bangladeshi immigrants in the UK are working and influencing policies in favor of their homeland, Bangladesh.

Moreover, with regard to identity maintenance, their activities in the host society expressed their commitment to retain their Bangla identity by maintaining strong linkages with their home culture. They present their vibrant and distinctive *Banglaness* by displaying their national cultural diversity in their host societal environment. They organize different cultural events, sports events, and celebrate all of Bangladesh's national days and festivals. Their distinctive Bangla identity is also expressed in the visible symbols of Bangladeshi monument, such as the *Shaheed Minar* and the sculpture of the Bangladeshi national flower *Shapla* that have been built by the collective initiative of the Bangla community in the UK. Therefore their activities show that they are working like a diaspora as the three broad elements of diaspora were clearly evident in our above findings and discussion.

References

Alexander, C., Firoz, S. and Rashid, N. (2003). *The Bengali Diaspora in Britain: A Review of the Literature*, London School of Economics, available at: http://www.banglastories.org/uploads/Literature_review.pdf [Accessed: 4 February 2012].

Brubaker, R. (2005). The 'diaspora' diaspora, *Ethnic and Racial Studies*, 28 (1), 1–19.

Change Institute (2009). *The Bangladeshi Muslim Community in England: Understanding Muslim Ethnic Communities*, Community and Local Government: London.

Castles, S. and Miller, M., J. (2009). *The Age of Migration: International Population Movements in the Modern World*. 4th edition. Basingstoke: Palgrave MacMillan.

Central Bank of Bangladesh, 2003. Economic Data: Wage Earners Remittance Inflows: Various tables. Dhaka: Government of Bangladesh.

Choi, I. (2003). Korean Diaspora in the making: Its current status and impact on the Korean Economy. In C.F. Bergsten and I. Choi (Eds.), *The Korean Diaspora in the World Economy* (pp. 9-29). Washington: Peterson Institute for International Economics.

Cohen, R. (1997). *Global Diaspora: An introduction*. London: UCL Press.

Cohen, R. (2008). *Global Diaspora: An introduction*. 2nd edition. London and New York: Routledge Taylor & Francis Group

Dufoix, S. (2007). *Diaspora*. Berkeley, CA and London: University of California Press

Eade, J and Garbin, D. (2003). *Bangladeshi Diaspora: Community Dynamics, Transnational Politics and Islamic Activities (with a focus on Tower*

Hamlets, Oldham and Birmingham), Report for the Foreign and Commonwealth Office.

Eade, J., & Garbin, D. (2006). Competing visions of identity and space: Bangladeshi Muslims in Britain. *Contemporary South Asia*, 15(2), 181-193.

Faist, T. (2010). Diaspora and Transnationalism: What kind of dance partners? In: R. Bauböck and T. Faist (Eds). *Diaspora and Transnationalism: Concepts, Theories and Methods* (pp. 9-34). Amsterdam: Amsterdam University Press.

Garbin, D. (2005). *Bangladeshi diaspora in the UK: some observations on socio-cultural dynamics, religious trends and transnational politics*. Paper presented at Conference on Human Rights and Bangladesh, Thornhaugh Street, London 17 June 2005.

Garbin, D. (2009). A Diasporic Sense of Place: Dynamics of Specialization and Transnational Political Fields among Bangladeshi Muslims in Britain, In M. P. Smith & J. Eade (Eds.), *Transnational ties: cities, identities, and migrations*. (Vol. 9, pp. 147-161). Transaction Publishers: New Brunswick (U.S.A.) and London (U.K.)

Gardner, K. (1993). Desh-Bidesh: Sylheti images of home and away. *Man*, 28(1), 1-15.

Gardner, K. (2006). The transnational work of kinship and caring: Bengali-British marriages in historical perspective. *Global Networks*, 6(4), 373-387.

Hussain, K., and Mirza, T. (2012). Bangladeshi Diet, In A. Thaker & A. Barton (Eds.), *Multicultural Handbook of Food, Nutrition and Dietetics* (pp.73-86). John Wiley & Sons.

Jennings, H. M., Thompson, J. L., Merrell, J., Bogin, B., & Heinrich, M. (2014). Food, home and health: the meanings of food amongst Bengali Women in London. *Journal of Ethnobiology and Ethnomedicine*, 10(1), 2-13.

Lawson, S., & Sachdev, I. (2004). Identity, Language Use, and Attitudes Some Sylheti-Bangladeshi Data from London, UK. *Journal of Language and Social Psychology*, 23(1), 49-69.

Mand, K. (2010). 'I've got two houses. One in Bangladesh and one in London... everybody has': Home, locality and belonging (s). *Childhood*, 17(2), 273-287.

Safran, W. (1991). Diasporas in Modern Societies: Myths of Homeland and Return. *Diaspora* 1(1), 83-99

Siddiqui, T. (2004). *Institutionalizing Diaspora Linkage: The Emigrant Bangladeshi in UK and USA*. RMMRU, Dhaka, available at: http://www.samren.net/Research_Papers/doc/Institutionalising%20Diaspora%20Linkage%20-%20The%20Emigrant%20Bangladeshis%20in%20UK%20and%20USA%20bangladesh%20diaspora.pdf [Accessed: 8 February 2012]

Shikder, J., U., M. (2008). Bangladesh. *Asian and Pacific Migration Journal*, 17(3-4), 257-275.

Verhulst S. (1999). Diasporic and Transnational Communication: Technologies, Policies and Regulation, *the public*, 6 (1), 29-36.

Vertovec, S. (1999). Three meanings of "diaspora," exemplified among South Asian religions. *Diaspora: A Journal of Transnational Studies*, 6(3), 277-299.

Wahlbeck Ö. (1998). *Transnationalism and Diaspora: The Kurdish Example*. Paper presented at 16th World Congress of Sociology, Montreal, Canada, 26 July – 1 August 1998.

Zeitlyn, B. (2011) Conceptualising Simultaneity in the British Bangladeshi Transnational Social Field. Paper presented at the International RC21 conference- The struggle to belong: Dealing with diversity in 21st century urban settings. Amsterdam, 7-9 July 2011

RSC, Number 6, Issue 2, May 2014, pp. 129-153.

Social change and economic transformations. Can variations in social capital across time affect the structure of an economy?

Johannes Zück

Bielefeld University, Germany
johannes.zueck@uni-bielefeld

Abstract: *Drawing on a body of literature that describes how the social capital affects economic life, this paper discusses the transformation of economies. Based on indicators of trust and norms of civic cooperation from the World Values Survey that are applied to a sample of 29-nation sample I present evidence that variations of social capital across the time actually can influence the structure of an economy. The research indicates a strong positive correlation between both trust and norms of civic cooperation and the creation of new firms that shape the economy.*

Keywords: Values, Social Capital, Trust, Civic Cooperation, Economic Structure, Social Change.

At least since the work done by Pierre Bourdieu (1986) the relationship between social capital and economic structures has been an important topic in economic sociological research. In this context, this paper will present evidence for a measurable impact of social capital on the structure of an economy. The findings of the empirical study are twofold: First it turns out that – compared between different states – variance in social capital results in a different economic structure. Furthermore the change of social capital over time has an own impact on structural changes of an economy: this research gives evidence that a growth of social capital over time translates into a denser economic structure.

On the way to this result, I, *first* conceptualized the concepts of economic structure and social capital. Literature has dealt with these topics in manifold ways. I will feature some of them in order to find proxies for measuring both. *Second*, these findings will be transferred into concrete variables. The economic structure will be measured by economic performance and the emergence of new business organizations. Social capital will be expressed as norms of civic cooperation and trust. *Third*, I will prove the correlation of both by an ordinary least squared regression analysis. Similar as in Knack & Keefer the regression analysis is based on a 29-nation sample. The huge difference is that this paper does not only focus on market economies. The sample therefore contains a huge diversity of economies to cope with the diversity of possible consequences of social capital for the structure of an economy. The influence of variations over time will be observed by a comparison between data from the sixth wave of the World Values Survey (WVS) from 2010-2014 and data from the fourth wave (1999-2004). This part of the paper also analyses the implications of social capital for the economic structure. I will conclude with a brief summary of the findings and a prospect for further research.

Theory – or, what matters?

The guiding question for this part is the search for an understanding of economic structures on a nation state level. I have two reasons for focussing on the nation state. First, this constraint enables actual measurability of the economy's properties – other than more holistic approaches.³ Second, the focus on national economies is also prevalent in existing accounts on the relationship between culture and structure (see for instance Coates 2000; Putnam 2000; Hall & Soskice 2001; or Amable 2003). But what is it that is described as economic structure? According to Chris Howell different theories on national economies are unified on their view of interdependent political-economic institutions: national capitalisms are characterized as “particular configurations of interlocking and interdependent political-economic institutions” (Howell 2003: 103). The “Varieties of Capitalism” approach by Peter A. Hall and David Soskice regards corporations as decisive economic actors (Hall & Soskice 2001). By following this approach I will emphasise on firms and how they coordinate with other actors.

Beside firms, there is another important aspect, which is part of the economic structure. Neil Fligstein introduces four institutions that “enable actors in markets to organize themselves, to compete and cooperate, and to exchange” (1996: 658). Only with these institutions – property rights, governance structures, conceptions of control, and rules of exchange – the structure of an economy can be defined. In other words the economic structure is described as the expectations created by institutions that have a

³ A famous representative of holistic analysis is Niklas Luhmann's theory on society (for the account on economy see Luhmann 1988). Boldyrev (2013) summarised Luhmann's contribution for economics.

high impact on the way firms operate in markets. The important task of this research is constructing a framework in order to quantitatively measure the structure of economies. As we just examined this measurement involves not only single firms and actors; it rather brings firm-influencing institutions into focus. A way to capture firm-influencing elements as well as the firms as a part of the structure lies in the measurement of economic performance. This needs some explanation: other than Knack & Keefer (1997) in their study on the impact of social capital on the economy I mainly focus on the structure, not the performance. But the performance can be an indicator for the structure. At the heart of neo-institutionalism lies the assumption that institutions rather than rational choice-driven actors structure the economy (DiMaggio & Powell 1983). As Douglass North (1990: 5) argues, “institutions affect the performance of the economy by their effect on the costs of exchange and production”. Thus, structural changes in the economy find expression in a changing economic performance. In other words, to pursue the question of this paper we should identify variance in economic performance. Before moving on to the next part I will briefly touch the subject of social capital.

Fligstein (1996) already describes the connection between culture respectively social capital and market institutions. This advises us to contrive an understanding of the term “social capital”. At the latest when it is about measuring social capital we must beware a squishy concept. So before explaining the proxies for economic structure and social capital I will briefly introduce the theoretical background of social capital. More than twenty years ago James Coleman described personal capabilities as “authority relations, relations of trust, and consensual allocation of rights which establish norms” (Coleman 1990: 300-1).⁴ This view on individual capability can be used as a definition of social capital (Knack & Keefer 1997: 1252). Using this definition for social capital has two advantages, first, it makes this research comparable

⁴ Due to constraints another important approach by Pierre Bourdieu (1986) can only be mentioned.

to the paper by Knack & Keefer (1997), as well as it refers thereby to an already accepted method of measurement for social capital. Numerous studies on the relationship between social capital and economy in the wider sense resulted in fruitful insights. Just to name Greif (1989) who exemplified how the development of trust as social capital in the Middle Ages influenced trade. Or, the paper by Helliwell & Putnam (1995), which showed higher growth rates in regions with a more developed “civic community”. Therefore, in our results we should also find this positive connection between social capital and the economic structure. After laying the groundwork for an understanding of social capital and the structure of an economy we will move forward to the question how we can actually measure these phenomena.

Data and methods – or, what to measure?

This chapter will determine what kind of things we have to measure in order to measure “the right things” – described above as social capital and economic structure. We need quantifiable parameters that already have been measured. In other words, I have to identify characteristics of social capital as well as of economic structures that can be found in a dataset. Another important criterion for finding appropriate proxies is the matter of comparability. This paper wants to make a statement about the impact of variations of social capital over time. To do so I will have to compare the findings with the results of Knack & Keefer (1997). Thereby I will also follow their definition of social capital by norms of civic cooperation and trust. Let’s have a closer look on how they fit this our understanding of social capital and their possible connection to the economic structure. Civic norms are defined as “those that resolve prisoner’s dilemmas without imposing substantial external costs on other parties” (Knack & Keefer 1997: 1254). In other words, these norms contribute to decisions that, but a common interest in the foreground rather than pure self-interest – albeit the aggregated consequences of such

“cooperative decisions” fit the personal interest best.⁵ Since civic norms constrain opportunism a variance of such norms is supposed to have an impact on economic structure. Lower costs to monitor and enforce contracts, less patent lawsuits, or more inter-organisational cooperation may result from high values of civic cooperation. So norms of civic cooperation are not only about the question if one puts his or her chewing gum in the next bin or on the street, it has the power to change the economic structure.

Trust also is a vital proxy for social capital and offers a link to economic structure. Putnam et al. (1993) utilise trust as one important brick in their framework of social capital that serves as his tool for analysing the governmental and economic capacity in Italy. They observed that a well working economy and well-functioning political system were the consequence of high values of social capital. Likewise James Coleman stresses the interrelation between trust and economy: “norms, interpersonal trust, social networks, and social organization are important in the functioning not only of the society but also of the economy” (Coleman 1988: 96). So it seems that trust, as well as civic cooperation, is a suitable concept for measuring social capital.

To measure trust and civic cooperation I will use questions from the sixth wave of the World Values Survey that has been collected from 2010 till 2014 (WVS 2014). Although I try to follow Knack & Keefer’s concept, I have to apply small modifications to the measurement of civic cooperation. Knack & Keefer (1997) took the responses to the question if each of the following behaviours “can always be justified, never be justified, or something in between” (1256):

⁵ A famous example for the benefit of cooperative decisions is the “tragedy of commons” (Hardin 1968).

- a. "Claiming government benefits which you are not entitled to"
- b. "Avoiding a fare on public transport"
- c. "Cheating on taxes if you have the chance"
- d. "Keeping money that you have found"
- e. "Failing to report damage you've done accidentally to a parked vehicle"

Since the items "keeping money that you have found" and "failing to report damage you've done accidentally to a parked vehicle" are not part of the 2014 WVS anymore I have to drop them out. Instead I will add the item "someone accepting a bribe in the course of their duties" as another indicator for the level of civic cooperation (CIVIC). The responses to each of those items – on a range between 1 and 10 – will be aggregated to a new item with a range of 4 to 40 (Knack & Keefer 1997: 1257). In our case 40 represents a high level of cooperation.

The question used for measuring trust (TRUST) is extracted from the World Values Survey: "Generally speaking, would you say that most people can be trusted, or that you can't be too careful in dealing with people?" (WVS 2014). This item embraces only a two option answer: "Most people can be trusted" or "Need to be very careful". After recoding the average value of this variable it represents for each country the rate of respondents that think most people can be trusted.

The proxies for the economic structure are a little harder to find. There is no straightforward way to measure market institutions like property rights, governance structures, conceptions of control, and rules of exchange as well as firm interrelations. We have to find a quantitative, measurable proxy for qualitative properties.⁶ According to the outlined framework I argue that

⁶ This is called "commensuration" as it is characterised as "the transformation of qualities into quantities that share a metric, a process that is fundamental to measurement" (Espeland & Sauder 2007: 16).

economic performance – measured by growth and the number of founded firms – is part of the economic structure. Measuring change in the economic performance also facilitates implications on change in the economic structure. Therefore, this paper utilises two indicators: (1) as well as Knack & Keefer (1997) I will use average annual growth per capita.⁷ Furthermore, I choose (2) the average number of newly registered firms per working-age people from 2004 till 2012 as an indicator of economic structure.⁸ It is obvious that both numbers – even in combination – do not give a meaningful number for the structure of an economy at a particular point in time. But that does not matter since the proxy is only important to identify relative changes over the years, to answer the question about the impact of social capital.⁹

I will answer the question by a bifid process. *First*, I will use the items exposed above to calculate regression models that give evidence about the relation between social capital and the economic structure. In this case an additional explanatory variable will complement all models: the individual perception of science enhances the model (SCIENCE). In the fashion of CIVIC this variable is based on the answers of the question “how much you agree or disagree with each of these statements” (WVS 2014):

- a. “Science and technology are making our lives healthier, easier, and more comfortable”

⁷ 2004-2012; World Development Indicators (<http://data.worldbank.org>)

⁸ The data is based on the reports of company registrars on the number of new firms. “Business entry density is defined as the number of newly registered corporations per 1,000 working-age people (those ages 15–64).” (<http://www.doingbusiness.org/data/exploretopics/entrepreneurship>)

⁹ Labelling economies at a certain point with a single number is what rating agencies are doing.

- b. “Because of science and technology, there will be more opportunities for the next generation”
- c. “We depend too much on science and not enough on faith”

Furthermore, the response to the question “All things considered, would you say that the world is better off, or worse off, because of science and technology?” is taken into account. The response of all four items – on a range between 1 and 10 – will be aggregated and into a new variable SCIENCE with the range from 4 to 40.¹⁰ As well as for CIVIC 4 equals a low affection to science, 40 a high one. The *second* step is a comparison between the results of these regression models – based on the 2014 WVS – with data from the fourth wave of the WVS (1999-2004). This will finally allow answering the question if variations in social capital across time can affect the structure of an economy.

¹⁰ Certainly the response to the statement “we depend too much on science and not enough on faith” will be recoded, to fit the others statements trends.

Table 1: Results – or, how social capital affects the economic structure¹¹

Can variations in social capital affect the structure of an economy? (OLS)				
Equation	1	3	4	6
Dependent variable	Growth	Growth	New firms	New firms
Constant	-7,559 (5,154)	-6,476 (6.494)	** 21.611 (9.14)	14,394 (11.272)
SCIENCE	** 0.441 (0,19)	** 0.447 (0.194)	** -0.803 (0.336)	** -0.841 (0.337)
TRUST	*** -0.0615 (0.02)	*** -0.061 (0.02)	*** 0.165 (0.035)	*** 0.16 (0.035)
CIVIC		-0,0361 (0.128)		0,241 (0.222)
Adj. R ²	0,24	0,21	0,42	0,42
	* p < 0.1	** p < 0.05	*** p < 0.01	

First, we have to find evidence of any relation between social capital and the economic structure, before we can proceed further and draw our attention towards the more interesting case of variations of social capital over time. Six regression models are the result of the first part of the empirical research, but only four of them possess a significant F-value.¹² Therefore, Table 1 only shows the significant regression models – without the equations 2 and 5. In the following lines I will explain the findings of the four regression models, before continuing towards the results for variances in social capital over time.

The dependent variable in equations 1 and 3 of Table 1 is average annual growth per capita in percent over the 2004-2012 period. In both equations the variables TRUST and SCIENCE show a strong relationship to

¹¹ Standard errors are in parentheses.

¹² All models are in the appendix.

variables of the economic structure. Interesting is the role of CIVIC. The variable has no significant relation to economic growth and as the values for the adjusted R^2 show it even weakens the explanatory power of the regression model. Therefore, I will leave equation 3 out of the analysis and focus on equation 1. The coefficient for SCIENCE suggests that a one-point rise in the 4 to 40 scale of science affection is associated with a 0.44 rise in the percentage of average annual growth. A ten percentage-point rise of TRUST is linked with a fall of annual growth by 0.62 percentage points. While the former – the strong relation between SCIENCE and growth – seems to be comprehensible this turns out to be a very interesting fact. One reason for this negative correlation may be that trust is higher in countries that already feature a strong economy. So the relative growth may be lower compared to countries with lower trust and a lower level of economy. Since TRUST is highest in the Netherlands and Sweden the suspicion arises that also the Euro crisis plays a role in this result: compared to the rest of the country sample Sweden and the Netherlands show a relatively low rate of average annual growth of 1.51% respectively 0.86%. This contributes largely to the negative correlation in the equations 1 and 3.

Equations 4 and 6 show the relation between the average number of newly funded firms per 1.000 people in working age and SCIENCE, TRUST and CIVIC. As well as in the first models CIVIC is not significant. Since model 6 with CIVIC also does not entail a bigger explanatory value I will focus the analysis on equation 4. Interesting is that the explanatory value of this model (adj. $R^2 = 0.42$) is far higher compared to equation 1. SCIENCE and TRUST are better to explain the number of newly founded businesses than economic growth. Precise a five-point rise in SCIENCE decreases the annual average number of new businesses per working-age people by four. A ten-percentage-point rise of TRUST leads to 1.6 more businesses per working-age people. Therefore the results exhibit a strong positive relation between TRUST and new firms as

indicator of the economic structure.¹³ This result perfectly fits the findings of Putnam et al. (1993) and Coleman on the relevance of trust for the economic structure. In a nutshell, this first step of analysis revealed that there is a relationship between social capital and the economic structure. Even if the trend of this relation in case of TRUST and economic growth does not fit the assumptions the positive correlation between TRUST and the number of new businesses elucidates the importance of trust as social capital for the structure of an economy. In the next step I will show the findings on how changes over time affect the economic structure.

Table 2.¹⁴

Can variations in social capital across time affect the structure of an economy? (OLS)	
Equation	6
Dependent variable	New firms
Constant	*** 3.612 (.557)
TRUST	*** .172 (.039)
CIVIC	*** .754 (.208)
Adj. R ²	0,69
	* p < 0.1 ** p < 0.05 *** p < 0.01

To examine the effects of variations of social capital across time I used data from the fourth wave of the WVS and compared it with results from the

¹³ The relation between SCIENCE and new firms is even stronger, but negative. Since SCIENCE is not in the focus of the research question, I will leave further interpretations out.

¹⁴ Standard errors are in parentheses.

current wave. It was possible to create the same variables for TRUST and CIVIC since the items have been the same. The final regression models tested the correlation of growth and new businesses with the difference of trust in a ten-country sample between 2004 and 2014.¹⁵ Again six models have been calculated, only one of them turned out to be significant. In this equation the difference of TRUST and CIVIC have a positive effect on the emergence of new businesses: if the value of trust in the 2004 survey is increased by six the annual number of newly founded firms per working-age individuals raise by one between 2004 and 2012. A four-point rise in the scale of CIVIC from 2004 till 2014 is associated with a three point increases of the “new firms” index. This model, however, explains a huge part of the number of new businesses with an adjusted R^2 of 0.69. To summarise: the regression analysis between the number of new firms as response variable and changes of trust and civic cooperation as explanatory variables show a strong positive correlation between variations of social capital across time and the economic structure. This correlation is highly significant.

Conclusion

This research did reveal the relation between social capital and the structure of an economy. Based on a 29-nation sample I investigated this correlation in a two-stage process. First I have drawn attention to the general impact of social capital on the economic structure. In this part of the analysis a relationship between trust and the economic structure has been discovered. Whereas the positive effect of trust on the emergence of new firms has been expected by theory the negative impact on economic growth was peculiar. This odd correlation may be explained by not included side effects that have a bigger impact on the response variable than the response variables. In the case of trust and growth the Euro crisis may have adulterated the model. European

¹⁵ The sample size is restricted to ten due to the fact that only ten countries of the original 29-nation sample were part of the fourth WVS wave.

countries with high trust that used to have a well-performing economy suffered from the consequences of the economic crisis. Therefore the values of annual growth have been much lower compared to other countries in the sample. Consequently, it is not surprising that these results differ from the findings of Knack & Keefer (1997) that identified a positive correlation between social capital and the economic performance.

The big finding of the second part was that variations of social capital across the time actually can influence the structure of an economy. There is a strong positive correlation between both trust and norms of civic cooperation and the creation of new firms that shape the economy. If social capital increases over time this also has a positive impact on the economy. Not necessarily for the pure performance – since this regression model turned out to be not significant – but on business organisations that represent the structure. Compared with the literature this finding is not surprising on the basis of the general positive correlation between social capital and the economy. But it gives specific evidence for the positive correlation between a rise of trust and norms of civic cooperation over time, and a change of the structure of national economies. And for that reason it is still remarkable. Further research may put more emphasis on further factors that shape the impact of social capital on the economy. For example the existing economic structure might be relevant: to which degree does it matter if a liberal market economy or a coordinated market economy is observed?

References

Amable, Bruno (2003): *The Diversity of Modern Capitalism*. Oxford, New York: Oxford University Press.

Boldyrev, Ivan A. (2013): *Economy as a Social System: Niklas Luhmann's Contribution and its Significance for Economics*. In *American Journal of Economics and Sociology* 72 (2), pp. 265–292.

Bourdieu, Pierre (1986): *The Forms of Capital*. In John G. Richardson (Ed.): *Handbook of theory and research for the sociology of education*. New York: Greenwood Press, pp. 241–258.

Coates, David (2000): *Models of Capitalism. Growth and stagnation in the modern era*. Cambridge, MA: Polity Press; Blackwell.

Coleman, James S. (1988): *Social Capital in the Creation of Human Capital*. In *American Journal of Sociology* 94 (1), pp. 95–120.

Coleman, James S. (1990): *Foundations of Social theory*. 1st ed. Cambridge, MA: Harvard Univ. Press.

DiMaggio, Paul J.; Powell, Walter (1983): *The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields*. In *American Sociological Review* 48 (2), pp. 147–160.

Espeland, Wendy Nelson; Sauder, Michael (2007): *Rankings and Reactivity: How Public Measures Recreate Social Worlds*. In *American Journal of Sociology* 113 (1), pp. 1–40.

Fligstein, Neil (1996): *Markets as Politics: A Political-Cultural Approach to Market Institutions*. In *American Sociological Review* 61 (4), pp. 656–673.

Greif, Avner (1989): *Reputation and Coalitions in Medieval Trade: Maghribi Traders*. In *Journal of Economic History* 49 (4), pp. 857–882.

Hall, Peter A.; Soskice, David (2001): *Varieties of Capitalism. The institutional foundations of comparative advantage*. Oxford, New York: Oxford University Press.

Hampton, Philip (2005): *Reducing administrative burdens. Effective inspection and enforcement*. Norwich: HMSO.

Hardin, Garrett (1968): *The Tragedy of the Commons*. In *Science* 162 (3859), pp. 1243–1248.

Helliwell, John; Putnam, Robert (1995): *Economic Growth and Social Capital in Italy*. In *Eastern Economic Journal* 21 (3), pp. 295–307.

Howell, Chris (2003): *Varieties of Capitalism: And Then There Was One?* In *Comparative Politics* 36 (1), pp. 103–124.

Knack, Stephen; Keefer, Philip (1997): *Does Social Capital Have an Economic Payoff? A Cross-Country Investigation*. In *The Quarterly Journal of Economics* 112 (4), pp. 1251–1288.

Luhmann, Niklas (1988): *Die Wirtschaft der Gesellschaft*. 1st ed. Frankfurt am Main: Suhrkamp (Suhrkamp Taschenbuch Wissenschaft, 1152).

North, Douglass C. (1990): *Institutions, Institutional Change and Economic Performance*. Cambridge, New York: Cambridge University Press (The Political economy of institutions and decisions).

Putnam, Robert (2000): *Bowling Alone*. Paperback.: Simon & Schuster.

Putnam, Robert D.; Leonardi, Robert; Nanetti, Raffaella (1993): *Making democracy work. Civic traditions in modern Italy*. Princeton, N.J: Princeton University Press.

WORLD VALUES SURVEY Wave 4 1999-2004 OFFICIAL AGGREGATE v.20140429.
World Values Survey Association (www.worldvaluessurvey.org). Aggregate File
Producer: Asep/JDS, Madrid SPAIN

WORLD VALUES SURVEY Wave 6 2010-2014 OFFICIAL AGGREGATE v.20140429.
World Values Survey Association (www.worldvaluessurvey.org). Aggregate File
Producer: Asep/JDS, Madrid SPAIN

Data Appendix

Country	Trust 2014	Trust 2002	d_Trust	Civic 2014	Civic 2002	d_Civic	Science 2014	An. Growth*	N.B. Density*
Algeria	17,93	11,22	6,71	29,09	35,16	-6,06	27,87	1,42	0,44
Armenia	10,14			36,91			27,96	6,73	1,39
Australia	54,43			37,00			29,98	1,55	10,31
Belarus	35,17			32,28			29,39	7,60	0,66
Chile	12,77	23,01	-10,24	34,77	33,11	1,66	24,77	3,49	3,55
Colombia	4,13			34,53			24,25	3,34	1,41
Ghana	4,96			36,94			27,40	4,90	0,85
Japan	38,76	43,06	-4,30	37,87	37,35	0,53	28,72	0,79	1,09
Jordan	13,25	27,65	-14,40	37,07	38,44	-1,37	27,38	3,57	0,69
Kazakhstan	38,80			33,07			30,95	5,63	1,81
Malaysia	8,54			33,54			28,57	3,19	2,33
Mexico	12,42	21,84	-9,42	30,88	32,06	-1,18	25,75	1,51	0,73
Netherlands	67,42			37,47			29,00	0,86	4,49
New Zealand	56,78			36,61			26,97	0,86	19,66
Nigeria	14,78	25,59	-10,81	34,92	35,80	-0,88	26,76	4,22	0,65
Pakistan	23,92	30,83	-6,92	36,57	38,89	-2,32	29,49	2,47	0,04
Philippines	2,84	8,61	-5,76	27,49	30,40	-2,91	25,99	3,36	0,24
Romania	7,12			37,18			27,43	4,30	5,25
Russia	29,23			32,06			29,05	4,53	4,47
Rwanda	16,63			36,28			31,01	5,41	0,40
Singapore	38,52	14,71	23,81	32,84	34,98	-2,13	27,23	3,32	6,76
Slovenia	20,11			35,90			29,28	1,47	3,51
Spain	19,51	34,02	-14,51	36,58	35,90	0,68	26,61	-0,09	3,49
Sweden	64,85			35,16			31,27	1,51	4,80
Tunisia	16,00			35,69			28,21	2,74	1,15
Turkey	12,43			37,94			28,66	3,70	1,01
Ukraine	24,95			33,23			28,70	3,56	0,94
Uruguay	15,25			36,54			25,89	5,53	3,18
Uzbekistan	14,09			35,35			31,97	6,36	0,57
Average	23,99	24,05		34,89	35,21		28,15	2,96	3,37

*Annual average 2004-2012

Regression Models

*Economic structure and social capital

. regress var2 var5 var3

Source	SS	df	MS	Number of obs = 29		
-----+-----				F(2, 26) = 5.36		
Model	31.4418421	2	15.720921	Prob > F = 0.0112		
Residual	76.2036148	26	2.93090826	R-squared = 0.2921		
-----+-----				Adj R-squared = 0.2376		
Total	107.645457	28	3.8444806	Root MSE = 1.712		

var2	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
-----+-----						
var5	.4407113	.1895693	2.32	0.028	.0510461	.8303766
var3	-.061527	.019863	-3.10	0.005	-.102356	-.0206981
_cons	-7.558931	5.154209	-1.47	0.154	-18.15356	3.035697

. regress var2 var5 var4

Source	SS	df	MS	Number of obs = 29		
-----+-----				F(2, 26) = 0.57		
Model	4.55541649	2	2.27770825	Prob > F = 0.5700		
Residual	103.09004	26	3.96500155	R-squared = 0.0423		
-----+-----				Adj R-squared = -0.0313		
Total	107.645457	28	3.8444806	Root MSE = 1.9912		

var2	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
-----+-----						
var5	.2002921	.2007565	1.00	0.328	-.2123688	.612953
var4	-.0809489	.1450131	-0.56	0.581	-.3790276	.2171298
_cons	.5582884	6.914219	0.08	0.936	-13.65409	14.77067

. regress var2 var5 var3 var4

Source	SS	df	MS	Number of obs = 29		
-----+-----				F(3, 25) = 3.48		
Model	31.6854836	3	10.5618279	Prob > F = 0.0309		
Residual	75.9599732	25	3.03839893	R-squared = 0.2944		
-----+-----				Adj R-squared = 0.2097		
Total	107.645457	28	3.8444806	Root MSE = 1.7431		

```
-----
```

var2	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
var5	.4465115	.194098	2.30	0.030	.0467592	.8462638
var3	-.0608514	.0203642	-2.99	0.006	-.1027923	-.0189105
var4	-.0361962	.1278232	-0.28	0.779	-.299453	.2270606
_cons	-6.475625	6.49425	-1.00	0.328	-19.85078	6.899534

```
-----
```

. regress var1 var5 var3

```
-----
```

Source	SS	df	MS	Number of obs = 29		
				F(2, 26) = 11.01		
Model	202.928341	2	101.46417	Prob > F = 0.0003		
Residual	239.642331	26	9.21701272	R-squared = 0.4585		
				Adj R-squared = 0.4169		
Total	442.570671	28	15.8060954	Root MSE = 3.036		

```
-----
```

```
-----
```

var1	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
var5	-.8028086	.3361723	-2.39	0.024	-1.493821	-.1117965
var3	.1648045	.035224	4.68	0.000	.0924005	.2372084
_cons	21.61075	9.140206	2.36	0.026	2.822792	40.39872

```
-----
```

. regress var1 var5 var4

```
-----
```

Source	SS	df	MS	Number of obs = 29		
				F(2, 26) = 0.79		
Model	25.4633433	2	12.7316716	Prob > F = 0.4629		
Residual	417.107328	26	16.0425895	R-squared = 0.0575		
				Adj R-squared = -0.0150		
Total	442.570671	28	15.8060954	Root MSE = 4.0053		

```
-----
```

```
-----
```

var1	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
var5	-.1928193	.4038176	-0.48	0.637	-1.022878	.6372397
var4	.3590136	.2916909	1.23	0.229	-.2405656	.9585928
_cons	-4.135404	13.90781	-0.30	0.769	-32.72331	24.45251

```
-----
```

. regress var1 var5 var3 var4

```
-----
```

Source	SS	df	MS	Number of obs = 29		
--------	----	----	----	--------------------	--	--

```

-----+-----
                                F( 3, 25) = 7.78
    Model | 213.739942  3 71.2466472    Prob > F  = 0.0008
    Residual | 228.83073 25 9.15322919    R-squared  = 0.4830
-----+-----
                                Adj R-squared = 0.4209
    Total | 442.570671 28 15.8060954    Root MSE   = 3.0254

```

```

-----+-----
var1 |   Coef.  Std. Err.   t  P>|t|  [95% Conf. Interval]
-----+-----
var5 | -0.8414463  .3368882  -2.50  0.019  -1.53528  -0.1476121
var3 |  .1603036  .0353453   4.54  0.000   .0875085  .2330987
var4 |  .2411195  .2218577   1.09  0.287  -0.215805  .6980439
_cons | 14.39435 11.27181  1.28  0.213  -8.820381  37.60909
-----+-----

```

*Economic structure and variations of social capital over time
 . regress var1 var3

```

Source |   SS   df    MS       Number of obs = 10
-----+-----
                                F( 1, 8) = 0.06
    Model | .14478549  1 .14478549    Prob > F  = 0.8064
    Residual | 18.0418533  8 2.25523167    R-squared  = 0.0080
-----+-----
                                Adj R-squared = -0.1160
    Total | 18.1866388  9 2.02073765    Root MSE   = 1.5017

```

```

-----+-----
var1 |   Coef.  Std. Err.   t  P>|t|  [95% Conf. Interval]
-----+-----
var3 |  .0108508  .0428246   0.25  0.806  -0.087903  .1096045
_cons | 2.45574  .5138677   4.78  0.001  1.270759  3.640721
-----+-----

```

. regress var1 var4

```

Source |   SS   df    MS       Number of obs = 10
-----+-----
                                F( 1, 8) = 0.03
    Model | .073853636  1 .073853636    Prob > F  = 0.8612
    Residual | 18.1127852  8 2.26409815    R-squared  = 0.0041
-----+-----
                                Adj R-squared = -0.1204
    Total | 18.1866388  9 2.02073765    Root MSE   = 1.5047

```

```

-----+-----
var1 |   Coef.  Std. Err.   t  P>|t|  [95% Conf. Interval]
-----+-----
var4 | -0.0413904  .2291716  -0.18  0.861  -0.5698611  .4870803
_cons | 2.348136  .5736326   4.09  0.003  1.025337  3.670935
-----+-----

```

 . regress var1 var3 var4

Source	SS	df	MS	Number of obs = 10	
				F(2, 7) = 0.03	
Model	.154806998	2	.077403499	Prob > F = 0.9705	
Residual	18.0318318	7	2.57597598	R-squared = 0.0085	
				Adj R-squared = -0.2748	
Total	18.1866388	9	2.02073765	Root MSE = 1.605	

var1	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
var3	.0092719	.0523025	0.18	0.864	-.1144038	.1329475
var4	-.0174234	.2793427	-0.06	0.952	-.6779639	.6431171
_cons	2.424144	.7471386	3.24	0.014	.6574423	4.190847

 . regress var2 var3

Source	SS	df	MS	Number of obs = 10	
				F(1, 8) = 3.69	
Model	13.2768184	1	13.2768184	Prob > F = 0.0911	
Residual	28.8175438	8	3.60219297	R-squared = 0.3154	
				Adj R-squared = 0.2298	
Total	42.0943621	9	4.67715135	Root MSE = 1.8979	

var2	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
var3	.103907	.054123	1.92	0.091	-.0209007	.2287148
_cons	2.24431	.6494403	3.46	0.009	.7466979	3.741922

 . regress var2 var4

Source	SS	df	MS	Number of obs = 10	
				F(1, 8) = 0.87	
Model	4.11897798	1	4.11897798	Prob > F = 0.3788	
Residual	37.9753842	8	4.74692302	R-squared = 0.0979	
				Adj R-squared = -0.0149	
Total	42.0943621	9	4.67715135	Root MSE = 2.1787	

```
-----
```

var2	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
var4	.3091064	.3318328	0.93	0.379	-.4561014	1.074314
_cons	2.200131	.8306007	2.65	0.029	.284762	4.115499

```
-----
```

. regress var2 var3 var4

```
-----
```

Source	SS	df	MS	Number of obs =	10
				F(2, 7) =	11.19
Model	32.0647357	2	16.0323678	Prob > F =	0.0066
Residual	10.0296264	7	1.43280378	R-squared =	0.7617
				Adj R-squared =	0.6937
Total	42.0943621	9	4.67715135	Root MSE =	1.197

```
-----
```

```
-----
```

var2	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
var3	.1722698	.0390072	4.42	0.003	.0800325	.2645071
var4	.754407	.2083338	3.62	0.008	.2617758	1.247038
_cons	3.612346	.5572161	6.48	0.000	2.294739	4.929952

```
-----
```

RSC, Number 6, Issue 2, May 2014, pp. 153-183.

Informational value of data about data in surveys: example of two web surveys

Marija Paladin

University of Ljubljana

marija.paladin@gmail.com

Abstract: *Considering data about data in survey research a potentially rich source of additional research information in theoretical part paper discusses definitions, categorizations, usefulness and dilemmas connected to data about data. In empirical part of paper we presented some metadata, auxiliary data and paradata gathered in two web surveys conducted on www.1ka.si.*

We analyzed potential differences between pre and post reminder respondents in which we included partially and fully completed questionnaires. We also analyzed time spent by respondent to answer full questionnaire or each page of questionnaire. In this case we analyzed only those questionnaires that were fully completed.

Beside differences between type (size) of organization we were also interested if pre and post reminder participation in web survey and time needed to answer full questionnaire or each page depends on some control variables (age, work experience, education, gender).

Keywords: paradata, data about data, reminder, online questionnaire, web survey

Introduction

Researcher has to be aware that when conducting (social) survey or research she/he can reach far beyond survey product data or said in another way – just answers to his questions. Potentially useful tool are also byproduct data. We refer with this to data made in different aspects or phases of survey and which are not researcher's primary concern but could be useful in further analysis. Most often used expression describing these data is paradata, although we prefer somewhat broader expression 'data about data'. Detailed categorization different types of 'data about data' will be done in further text, especially considering paradata, metadata and auxiliary data.

Paradata is part of Couper's triple categorization of data, metadata, paradata (Kreuter, Couper & Lyberg, 2010: 286) and could be regarded as by-product of the (field, internet, phone) data collection process. Researchers use computer assisted methods to collect survey data and data about processes which allow statistical evaluation, monitoring and managing of survey process (Kreuter, Couper & Lyberg, 2010: 282).

Use of paradata focuses on management of some nowadays challenges such as declining response rates, increasing risk of non-response bias and measurement error, and escalating costs of survey data collection. The collection of survey paradata is not new but the range and detail of paradata being collected has increased due to the computerization of the survey process (Nicolaas, 2011: 4).

The possibilities of use of paradata for research purposes are wide and not so good explored. It is needed better knowing which paradata can be useful (and in connection to what), so that they will be collected and treated intentionally not just as byproduct. Some steps in that direction we are making in this paper.

This paper is divided into two parts, theoretical in which we talk about data which are produced in survey process (data about data) and empirical, which is based on two surveys conducted by an online questionnaire on www.1ka.si. Main focus and main goals were two. Firstly, to identify if there are some differences between those respondents who answered the questionnaire before or after reminder so we could tackle some information on pre or post reminder responders to address potential units of survey in invitation more appropriately (to get more responses). And second, to find out if there were some differences between respondents on the basis of time they needed to answer questionnaire completely to see if there exists any statistically significant differences between respondents based on how long they need to answer questionnaire.

Definition

Maybe because there is currently no consensus over a standard definition for paradata (Nicolaas, 2011) there are quite a few, more or less similar operational definitions of paradata. For example: Paradata are data collected about the survey process and captured during computer assisted data collection by interviewer's assistance or automatically. They include call records, interviewer observations, time stamps, keystroke data, travel and expense information, and other data (Kreuter, Couper & Lyberg, 2010: 282)

Other says that paradata are process data, or all the data collected during the response process and do not include the response itself and which exist in both interviewer-administered surveys and computer assisted self-administered surveys (Horwitz et. al. 2012).

Couper was the first author to introduce the term "paradata" to the field of survey methodology in terms of automatically generated process data. Now the term paradata covers all types of data about the process of collecting survey data such as interviewer call records, length of interview, interviewer characteristics, interviewer observations (Nicolaas, 2011: 3). Paradata are data captured throughout the entire survey process that are a result of collecting

the product data which could (or not) be used and intentionally collected (Frost & Duffey, 2010: 14)

Categorization

Four primary categories of survey data proposed by Frost Hubbard and Ben Duffey (2010: 15-16) and quite similar from Garry Nicolaas (2011: 3) are product data, paradata, metadata and auxiliary data.

Product data answers to questions in surveys or survey questionnaire data (not paradata). Paradata (process data) measure keystroke files, Contact attempt data, Interviewer hours worked and miles traveled, ect. Metadata describes variables, description of survey purpose ect. Metadata are static descriptions of a data file or data system, for example variable and value labels, response rates. Auxiliary data give information on sampling frames, Census area, characteristics, administrative data ect. Auxiliary data are pre-existing data that is used to support the survey process or the analysis of the substantive data.

For interviewer-administered surveys, paradata can include response times, respondent utterances (pauses, hedges, stutters), respondent expressions, interviewer observations. In computer assisted self-administered surveys, such as Internet surveys we can collect information about location of break offs, changed answers, error messages, mouse clicks, response times (Horwitz et. al. 2012: 1).

Paradata items that could be collected in computer-assisted personal interview surveys (Nicolaas, 2011: 7-12): Interviewer characteristics, call record data, interviewer observations about the area and dwelling, doorstep interaction, audit trails, audio recordings, other paradata items (data items which describe the process of asking and answering questions).

As we see some paradata are automatically collected (for ex. by software) and some are interviewer assisted (for ex. their observations). Different type of data collecting and also different type of data (time needed to complete questionnaire vs. observation on willingness to participate in research) may produce different quality of data. Despite that, as Casas-Cordero, Kreuter, Wang and Babey observed, the literature on the quality of, for example, neighbourhood observational data collected by interviewers is only now emerging. And due to moderate to low Cohen k statistics (which is used to score the agreement between observers in categorical rating tasks) they got in their research on neighborhood observations, we could argue that much of observed could have less relationship to the real characteristics of the areas than to characteristics of the interviewers (Casas-Cordero et. al., 2013: 228, 236-240).

The National Health Interview Survey (NHIS) uses following groups of paradata: response paradata, measures of time, measures of contactability, measures of cooperation, mode measures, survey-level Information (U.S. Department of Health and Human Services, 2012: 6-16).

Another categorization of paradata is that paradata can be macro or micro. As author says macro paradata or summary process measures is common and widely used. Examples of these are overall process summaries, like coverage rates, item and unit nonresponse rates. On the other hand there are micro paradata or process details known on each case, like language in which each interview was taken in multilingual environment, how many times the household was called before interviewing, whether there was refusal at the beginning ect. Micro paradata are less familiar. This could be due to lack of interest in seeing possible added value of paradata (by researchers) and possible additional expenses (on clients side). Micro paradata are not about overall survey process (aggregated) but rather it describes survey process on individual records (Scheuren: 1-2). We can understand macro paradata as metadata in previously mentioned categorization and micro paradata as paradata in previously mentioned categorization.

Collecting paradata is, beside in interviewer-administered surveys or laboratory-based research accessible in web surveys also. It is possible to conduct large-scale self-administered surveys while collecting paradata (Heerwegh a: 2).

Dirk Heerwegh, discussing mostly on audit trails, categorizes paradata in two subgroups: server side paradata and client side paradata. Mostly all web surveys gather some of server side paradata which are collected automatically without additional effort or consent of client (responder). If researcher is interested in deeper information on respondent behavior, like at the level of specific survey questions, author introduces term client side paradata. As author distinguishes, client side paradata are not collected at the level of the server, but at the level of the respondent's computer, by incorporated script (researcher detects respondent's behavior such as clicking radio-buttons, drop-boxes and hyperlinks) and are sent when responder submits the web page (Heerwegh a: 2-5). On this point we could argue that there are not client side paradata if there exists automatic collection, but on the other hand, some examples speak against it. For example, time needed for completing one page or all questionnaire, is on some survey software tools collected automatically when responder submits each page along with his/hers answers. Although submitting is needed, it's main purpose is to complete one block of questions and send answers. Submitting in this case does not mean that answers to survey questions are also paradata. But, never the less, we have to consider that author discusses about audit trail or audit log which is important for his categorization.

An interesting categorization offers Frauke Kreuter (2010: 4) when speaking about what paradata are available throughout the survey process. For this purpose she divides paradata into: key strokes, for example response times, vocal characteristics, for example pitch of interviewer voice, disfluencies, contact data and interviewer observation, for example day and time.

Categorization is especially interesting because it underlines one part of interviewers nonverbal communication (vocalic cues) as important part of survey process. Vocalic cues appeared to be quite important in many cases of face to face communication and also in process of persuasion which in other package takes place in convincing possible responder to participate in survey.

If we consider different categorizations properly we can make conclusion that to some degree there exist some overlapping between different categories. If we take categorization to product data, paradata (process data), metadata and auxiliary data, it must be said that there will be emphasis on all categories (without product data) for purposes of this paper, although primary interest lays in paradata.

Usefulness of Paradata

The possibilities of use of paradata in the spirit of statistical process control are wide. The first uses of paradata focused on exploration of measurement error in surveys. Paradata are also widely used to explore non-response in surveys, to manage data collection and for use of paradata-driven responsive design (Couper & Kreuter, 2013: 271).

Kreuter, Couper and Lybergz differ between paradata and their post-survey use and paradata used in monitoring and managing of ongoing surveys. *Paradata and their post-survey use* (Kreuter, Couper & Lybergz, 2010: 283-285) means post-survey assessments or post-survey corrections of errors common in the survey process. *Paradata in monitoring and managing of ongoing surveys* (Kreuter, Couper & Lybergz, 2010: 286-288) means that measures about the process are taken along the way, so that error sources can be located and interventions can be targeted during the collection process.

Paradata can be used to gain reliable and replicable findings about survey methods and practice to minimize survey error (Nicolaas, 2011: 4). They can also be used as alternative to measure to survey data quality analysis. Some argue that merely response rates are not most suitable measure of quality of survey data. That why it is proposed to upgrade this approach with other paradata. Pros of that point of view is that it involves more data, uses complete data, data are reported at the survey level, it encourages the development of paradata and cases are differentiated in process of (para)data collection. Approach is promising also because it enables comparison of respondents and nonrespondents on some variables. Further upgrading of this approach means comparison of respondents by paradata and data about crucial variables that is examining correlations between paradata and survey variables like comparison of early and late responders (Wagner 2009). Or to identify potential problems with the survey instrument, understand the process the respondent uses to complete the survey, to assess the quality of the instrument design, to evaluate how well the instrument is working and whether there are modifications that need to be made prior to production (Horwitz et. al. 2012).

Pro gathering and analyzing paradata reasons also arise from practice (Horwitz et. al., 2012). Among those reasons are identification of problematic screens or questions, testing usefulness of help option, identifying drop out points ect.

Dirk Heerwegh (a: 6-15) also sees a few possible uses of client side paradata, which can be categorized with following goals: calibrating progress indicators, testing the effects of response formats, testing the effects of question, identifying attitude strength.

Couper and Kreuter (2013) conducted exploratory study using paradata to explore item level response times in surveys on results of computer assisted survey from cycle 6 of the National Survey of FamilyGrowth (2002–2003). They found out that automatically derived indicators of item characteristics are found to vary systematically with response time and

interviewers also appear to contribute independently to the completion times (although it has to be stressed that measured demographic characteristics and experience of interviewers explain only a small part of variability) (Couper & Kreuter, 2013: 293-284).

SOME DILEMMAS ABOUT PARADATA

Paradata capture can be viewed as collecting information about the process of completing a survey. No behavior outside the survey is captured, so it can be argued that no additional consent than to participate in survey is needed, although the question of whether and how to inform respondents about the capture of paradata remains. On the other hand respondents are usually not aware that such additional information is being collected and, if they were aware of it that might change their behavior or decide not to participate in the survey. Questions is how to provide information about the collection of paradata, linkage them to survey data while at the same time maintaining respondent cooperation with the survey (Couper & Singer, 2013: 58-59).

Social surveys mostly rely on the voluntary cooperation of respondents and protection of their personal information and identities, actually and perceived. Some authors argue that various paradata include information that could disclose respondents' identities; for example address details, interviewer remarks, audio recordings. Consequently paradata databases cannot be released without thorough processing and the removal of information that could be used to identify respondents. But this process is problematic and time consuming (Nicolaas, 2011: 16).

Researchers must protect respondents from potential harm and assure their autonomy in deciding whether to participate in the research or not. This

means assuring and obtaining respondents' informed consent. To some authors this means assuring that they are treated as autonomous individuals with the right to make informed, voluntary decisions about participation. That is connected with ethical and practical questions arising from the growing use of paradata – the data collected by computerized systems during data collection – in surveys, especially those conducted online (Couper & Singer, 2013: 57).

Couper and Singer conducted a web study of how information about disclosure risk might affect survey participation. Results were following. 63.4% of those respondents who received a note describing a hypothetical survey and were then asked whether they would be willing to participate in the survey; if yes, whether they were willing to permit use of their paradata, agreed to do the survey and consented to paradata use. Same consent gave 59.2% respondents who received a note describing a hypothetical survey that they had already completed. Afterwards they were asked whether they would be willing to permit use of their paradata and 68.9% of them agreed. Differences between groups are statistically significant. Mentioning of paradata resulted in lower willingness to participate in the survey. Reasons respondents gave for refusing usage of their paradata were: concerns about aspects of paradata, with mentioning the tracking of browsing behavior, general privacy-related concerns. Many responses suggested confusion over the extent of paradata capture and additional explanation did not made things easier and more understandable (Couper & Singer, 2013: 63-65).

As authors said presented experiment did not adequately inform respondents about methodology around paradata and to elicit their consent. But on the other hand respondents are probably not aware that paradata are unavoidably collected in the process of responding to a survey so the question really is whether respondents would consent to their use or not (Couper & Singer, 2013: 65-66).

Heerwegh (a: 18) also opens the question about ethical concerns bonded to collection and analyzing paradata, because respondents may regard

it as a tool to invade their privacy. It could be understood that collecting client side paradata should occur only if it is the only way of answering a research question, and if it does not mean an invasion of privacy.

The question is also whether the use of paradata collected in web surveys reaches the level needing explicit mention to respondents. For many this arises ethical and legal dilemmas. Recent EU online privacy legislation and US regulations go in direction of requiring informed consent for the collection of any data other than the responses to the survey (Couper & Singer, 2013: 66).

With implementing of new Law on Electronic Communications in Slovenia in 2013 (<https://www.ip-rs.si/novice> ...), were brought new rules regarding the use of cookies and similar technologies for storing information or access to information stored on a computer or users mobile device. The new legislation does not prohibit the use of cookies, but exacerbates rules on conditions of how cookies and similar technologies may be used. The stress is given to the requirement that the users are paired and that they should be offered a choice of whether they allow or not websites to use cookies. The new legislation is primarily aimed at better protect of users' online privacy.

As can be seen from the Information Commissioner's guidelines on the use of cookies (<https://www.ip-rs.si/fileadmin> ...), that probable cookies originated in process of non-commercial research are not listed among the exceptions of cookies permitted for use without the prior consent nor among cookies which may not be used without previous consent of the user. However, it should be noted that the scope of usability of cookies is very vividly thus the guidance of the Information Commissioner will continue to be updated regularly.

EMIRICAL PART: PARADATA IN TWO WEB SURVEYS

The empirical part of this paper is based on two surveys conducted by an online questionnaire. The questionnaire was sent to Croatian small sized (up to 49 employees) and big sized (250 or more employees) organizations, with instructions to meet the person who is responsible for HRM and for the recruitment and employment of new staff in the organization. The target population (and sample in case of small organization) for questionnaire was determined with existing database in register of the Croatian Chamber of Commerce. Theme of the questionnaire (in both surveys identical) for the purpose of this paper was the impact of nonverbal factors on persuasiveness of individuals in business context. For purpose of this paper we are especially interested in paradata and some auxiliary data which were accessible on www.1ka.si and were gathered along with survey data.

Two main goals were, first to identify if there are some differences between those respondents who answered the questionnaire before or after reminder (there was only one reminder, which was sent to all units regardless of previous participation). Fully and partly completed questionnaires were included. And second, find out if there was some differences between respondents on the basis of time needed to answer questionnaire completely (only fully completed questionnaires included).

Table 1: Samples.

Type of org.	% HRM manager (referent positions excluded)	Age (mean)	Mean age of employ. (mean)	Work exper. in years (mean)	Duration of educat. in years (mean)	Number of employ. (mean)	Gender	
small (up to 49 empl.)	69%	40.9	37.6	20.4	16.0	64.1*	M	42 or 41%
							F	60 or 59%
big (from 250 empl.)	69%	39.0	39.6	14.9	16.6	439.4	M	34 or 39%
							F	54 or 61%

* Number exceeds 49, this could be due to variability in in data which is corrected in databases only in year interval, so some organizations included in sample exceed 49 employees. Also: not all respondents answered to question about organization.

DATA ABOUT DATA**Table 2: Sample frame.**

Type of organization	Population	Included in survey (sent invitation)	Planned response	Realized response (status 5 and 6)*
small (up to 49 employees)	75.917	1.333	Cca 10% eg. 130	7.88% eg. 105
big (from 250 employee)	449	414	Cca 10% eg. 40	21,50% eg. 89

* Partially full or completed questionnaire.

Table 3: Basic data about questionnaire and survey

Basic data about questionnaire and survey	Big organization	Small organization
Number of questions	17	17
Variables	91	91
Items	214	277
partially or completed questionnaires	89	105
Language	Hrvatski	Hrvatski
Estimated time for completion of q.	15min 0s	15min 0s
Real time respondent spent on q. (partially or complete)	<u>11min 54s</u>	<u>12min 43s</u>
Date of first item	6.1.2013	7.1.2013
Date of last item	28.1.2013	1.2.2013
Completed the survey (6)	69	89
Partially completed (5)	20	16
Total adequate (5+6)	89	105
Total inadequate	125	172
Total units	214	277

REMINDER AS STIMULUS TO PARTICIPATE**Table 4: Participation in survey before and after reminder.**

Type of organization	Invited to participate in survey	Participated	Participated before reminder	Participated after reminder
small (up to 49 employees)	1.333	105	32 (30,5 %)	73 (69,5 %)
big (from 250 employee)	414	89	13 (14,6 %)	76 (85,4 %)

Included: partially and fully completed questionnaires.

It is slightly surprising that we can see in table 6 that in small organization, where time pressure is maybe more important factor than in big organization, proportion of answered questionnaires before reminder was bigger (30,5%) in comparison to big organization (14,6%). In both cases though, most questionnaires were answered after reminder.

Table 5: Participation in survey before and after reminder - by gender.

Type of organization		Male	% (of male)	Female	% (of female)
small (up to 49 employees)	before reminder	14	33,3%	16	26,7%
	after reminder	28	66,7%	44	73,3%
big (from 250 employee)	before reminder	4	11,8%	9	16,7%
	after reminder	30	88,2%	45	83,3%

We see in table 5 that in small organizations male respondents in comparison to female respondents were slightly more willing to participate in survey before reminder. Right the opposite case was with respondents in big organizations. And in general, in both types of organization respondents of both genders were more willing to participate in survey after reminder.

In further we calculated percent of before and after reminder respondents on basis of position in organization and field of education. Due to space limitation we would not display all data in tables.

Analysis showed that in big organizations there exists the biggest percent of before reminder participants among respondents on owner position and smallest among respondents director of organization position. In small organizations there exists the biggest percent of before reminder participants among respondents on head of unit position and smallest among respondents director of district position. We have to note that the number of unit in some of groups is very small.

Analysis also showed that in big organizations there exists the biggest percent of before reminder participants among respondents with education from natural sciences and smallest among respondents with education in technical field. In small organizations there exists the biggest percent of before reminder participants again among respondents with education from natural sciences (the only case with more than 50% respondents from group in before reminder participation) and smallest among respondents in technical field and other. Like previously said, we have to note that the number of unit in some of groups is very small.

We also analyzed Pearson coefficient on duration of education, age and work experience in connection to size of organization to identify statistically significant differences. It showed that in our two surveys decision about participation in survey before one gets reminder does not depend

significantly on chosen demographic characteristics, with exception of age in case of respondents in small organization (in case of big organization no statistically significant differences showed).

TIME NEEDED TO ANSWER QUESTIONNAIRE

In this section there were analyzed only those questionnaires that were fully completed (partials are excluded). Excluded were also those questionnaires in which more than 45 minutes for completion were needed (more than 3 times exceeded estimated time). On that criterion 7 items from big organizations and 10 from small organizations were excluded.

The questionnaire, which was used in both discussed online surveys are composed of the following sets of questions divided into 6 pages.

Page 1 with demographic questions about the respondent, a set of seven statements about non-verbal communication of the respondent (5 Point Likert-type scales), a set of eleven statements about the factors of persuasion (5 Point Likert-type scales). Page 2 with set of 16 statements about factors of movement and touch (5 Point Likert-type scales). Page 3 and 4 with set of 30 statements about factors of appearance and decoration (5 Point Likert-type scales). Page 5 with set of 10 statements covering the vocalic factors (5 Point Likert-type scales) and a set of 3 Statements time factors (5 Point Likert-type scales). And page 6 with questions about the company.

From what we can see in Table 6 we have to argue that coefficients of Skewness and Kurtosis give us information about non-normal distribution in time needed to answer full or each page of questionnaire. That is why we also have to take into account medians which are in some cases very close to mean (2nd page about factors of movement and touch in both types of org., 3rd page about factors of appearance and decoration in small org., 4th page about

factors of appearance and decoration in both types, 5th page about vocalic and time factors in big org.) and quite different from mean in other cases (full questionnaire, 1st and 6th page in both types, 3rd page about factors of appearance and decoration in big org., 5th page about vocalic and time factors in small org.). Never the less we will in further analysis regard mean/average as appropriate measure of mean value (t-test, Pearson coefficients).

We can see that mean/average time needed for answering full questionnaire in both types of organization (calculated as average the difference in starting and ending time, pauses are not recorded and thus not taken into account) exceeds estimated time needed just slightly and that, on the other hand median is in both cases slightly under estimated time. As we also see in table 6, estimated time is exceeded in both types of organization in average time needed for answering to most of pages in comparison to estimated time. This appears regardless to page topic (discussed above). Exceptions are 2nd page about factors of movement and touch for both types, 3rd, 4th about factors of appearance and decoration and 5th covering the vocalic and time factors page in small organizations. We could argue that, due to longer time needed to really answer questions on each page in comparison to estimated time, that responder needed slightly more time to comprehend the topic of questions which was nonverbal communication cues based. These topics (especially for example appearance and touch) can be perceived as sensitive topics by many. That is why it is also useful information about median time needed for answering each page in comparison to estimated time. Due to non-normal distribution in time needed to answer each page of questionnaire. Median is slightly lower on almost all pages for both types of organizations than mean/average time (exception is 1st page for both types of organizations). This argues against some assumed difficulties in question comprehension mentioned above.

In further we calculated t-test for identifying possible statistically significant differences between two types of organization. It showed that there exists statistically significant difference only on average time needed to complete 5th page covering vocalic and time factors. On this page respondent in big organizations needed more time to complete the page than respondents in small organizations. In all other pages, including time needed to complete full questionnaire, existing real differences were not statistically significant (due to space limitation tables are not included in paper). We could argue that respondents from both types of organization had taken similar effort to answer questionnaire regardless to time pressure which is in small organizations is, presumed, to be higher.

We were also interested if time needed to answer full questionnaire or each page was statistically significantly connected with some control variables (age, work experience, education), if there exists some statistically significant differences in time based on gender of respondents. Firstly we wanted to know if there are some statistically significant differences between two types of organization on basis of chosen control variables. Computed t-test showed none statistically significant differences between respondents from big or small organizations on age, years of work experience and years of education (due to space limitation tables are not included in paper). Hi square test also showed no statistically significant differences between two types of organizations on gender of respondents (due to space limitation tables are not included in paper). Respondents from both types of organizations seems to be quite comparable due to chosen control variables.

In further we calculated Pearson's coefficients of correlation between time needed to answer full questionnaire or each page in questionnaire and previously mentioned control variables in both types of organizations (due to space limitation tables are not displayed). Some statistically significant correlations do exist and correlations are not the same if we consider type (size) of organization. Exception is 1st and 6th page where no significant correlations were calculated regardless to type of organization.

On one hand time needed to answer full questionnaire is in big organization statistically significantly correlated with two of three control variables (positive with age and work experience). On the other hand none statistically significant differences exists in case of small organizations.

If we consider each page separately, we see that in total more statistically significant correlations exists in case of big type of organization (8 in big type and just 2 in small type of organization). Also interestingly, time needed to answer to 2nd page about factors of movement and touch is in case of big organization statistically significantly correlated to all three chosen control variable (positive with age and work experience and negative with education) on one hand. And on the other, in case of small organization just one statistically significant correlation exists (positive with work experience).

Other statistically significant correlations are as follows. In small organizations: in time needed to answer to 3rd page about factors of appearance and decoration with work experience (positive). In big organization: in time needed to answer to 4th page about factors of appearance and decoration with education (negative) and in time needed to answer to 5th page covering vocalic and time factors with age (positive) and work experience (positive).

We also calculated t-test for identifying possible statistically significant differences in time needed to answer full questionnaire or each page in questionnaire regarding to gender of respondents in each type of organization. It showed that there exist two statistically significant differences in case of small organizations and one in case of big type of organization. In case of big organization male respondents in comparison to female respondents needed more time to answer 1st page about respondent's demographics, non-verbal communication of the respondent and about the factors of persuasion. In small type of organization male respondents in comparison to female respondents needed more time to answer 2nd page about factors of movement and touch. Female respondents in comparison to male respondents

also needed more time to answer 5th page covering vocalic and time factors (due to space limitation tables are not included in paper).

CONCLUSION

In theoretical part of paper we first discussed definitions on metadata, auxiliary data and paradata – data about data. Main focus was given to data about data in case of web surveys where nowadays interesting area audit trail is. Beside potentially fruitful role of gathering and analyzing paradata in (web) surveys, some concerns also arises. There are open some ethical dilemmas bonded to collection and analyzing paradata due to possible regarding paradata as a tool to invade responders privacy. But in opinion of some authors (and also ours), the real question one has to answer is whether the use of paradata collected in web surveys really reaches the confidentiality and other form of threat to respondents privacy to that point or level on which it is needed explicit mention. But on the other hand, with implementing of new Law on Electronic Communications in Slovenia in 2013 researcher have to be aware what legal and other connotation may be given also to area of paradata. In new law there were brought new rules regarding the use of cookies and similar technologies for storing information or access to information stored on a computer or users mobile device.

In empirical part of paper we presented some metadata, auxiliary data and paradata gathered in two web surveys conducted on www.1ka.si. It is interesting to underline some results. In part where we were analyzing potential differences between pre and post reminder respondents in which we included partially and fully completed questionnaires, some results were quite interesting. In small organization, for example, where time pressure is maybe more important factor than in big organization, proportion of answered questionnaires before reminder was bigger in comparison to big organization. Although in both cases, most questionnaires were answered after reminder. It also seems that in our two surveys decision about participation in survey before one gets reminder does in some cases depend on chosen demographic

characteristics. For example, in small organizations male respondents in comparison to female respondents were slightly more willing to participate in survey before reminder. Right the opposite case was with respondents in big organizations. In big and small organizations there exists the biggest percent of before reminder participants among respondents with education from natural sciences and smallest among respondents with education in technical field. In case of respondents in small organization before reminder respondents are older.

In part in which we were interested in time spent by respondent to answer full questionnaire or each page of questionnaire were analyzed only those questionnaires that were fully completed (partials are excluded). Excluded were also those questionnaires in which more than 45 minutes for completion were needed (more than 3 times exceeded estimated time). Mean/average time needed for answering full questionnaire in both types of organization exceeds estimated time just slightly and that, on the other hand median is in both cases slightly under estimated time. Estimated time is exceeded in both types of organization in average time needed for answering to most of pages in comparison to estimated time (regardless to topic of page). We could argue that, due to longer time needed to answer questions on each page in comparison to estimated time, that responder needed slightly more time than it was assumed to comprehend the topic of questions which was nonverbal communication cues based. These topics (especially for example appearance and touch) can be perceived as sensitive topics by many. But in the other hand information about median time needed for answering each page in comparison to estimated time (which is slightly lower on almost all pages for both types of organizations) argues against assumed difficulties in question comprehension mentioned above.

Calculated t-test for identifying possible statistically significant differences between two types of organization showed that just on one page (5th) respondent in big organizations needed more time to complete the page than respondents in small organizations. It could be argued that respondents

from both types of organization had taken similar effort to answer questionnaire regardless to time pressure which is in small organizations, presumed, to be higher.

We were also interested if time needed to answer full questionnaire or each page was statistically significantly connected with some control variables (age, work experience, education) and if there exists some statistically significant differences in time based on gender of respondents. Firstly we wanted to know if there are some statistically significant differences between two types of organization on basis of chosen control variables. Analysis showed that respondents from both types of organizations seemed to be quite comparable due to chosen control variables. Due to calculated Pearson's coefficients of correlation between times needed to answer full questionnaire or each page in questionnaire and previously mentioned control variables in both types of organizations some statistically significant correlations do exist. It is interesting that on one hand time needed to answer full questionnaire is in big organization statistically significantly and positively correlated with two of three control variables (age and work experience). On the other hand none statistically significant differences exist in case of small organizations. If we consider each page separately, we see that in total more statistically significant correlations exist in case of big type of organization (6 in big type and just 2 in small type of organization).

Calculated t-test for identifying possible statistically significant differences in time needed to answer full questionnaire or each page in questionnaire regarding to gender of respondents in each type of organization showed some significant differences. In case of big organization male respondents in comparison to female respondents needed more time to answer 1st page which comprised (beside questions about respondent's demographics and about the factors of persuasion) some questions about self-evaluation of respondent's non-verbal communication. In small type of organization male respondents in comparison to female respondents needed more time to answer 2nd page about factors of movement and touch. Previous

researches about nonverbal communication showed, that in many cases men have more trouble excepting touch as appropriate way of communication than women.

In case of presented two surveys we operated with limited number of paradata which showed to be potentially informative and useful although there are some limitations which will be noted in further. This goes hand in hand with examples of other authors who indicate that paradata seems to be very useful tool to evaluate several points in survey, survey data and survey process. As some authors say, more methodological research is required to identify the key paradata items to be collected and the best ways to use those (Nicolaas, 2011: 4). That is why the use of paradata is still in need for development. Many also argue that little is known about the quality of paradata and, consequently the usefulness of the data (Nicolaas, 2011: 17). There are also some dilemmas about problematic viewpoints of gathering and usage of paradata connected with confidentiality of respondent and also interviewer. In my opinion it has to be carefully identified where paradata opens real confidential and other ethical dilemmas and where it is simply too much emphasis giving to unreal threat.

There exists some of limitations that we have to highlight. Samples in both surveys are rather small to make definite conclusions in some cases. Although reminder was just one it was sent in two waves (1/2 respondents in first wave, 2/2 in second), so there could be some differences between first and second wave which were not analyzed (even though every respondent was sent just one reminder).

Data base for contacts of units invited to survey may not be updated completely (differences in number of employees, actual existence of organization due to economic crisis, accessibility of e-mail – not for all, ect.) due to one year interval for updating information, freedom of organizations to give some information, like e-mail ect.

It would be very useful to have more detailed paradata referring on each question of questionnaire. This could give more in-depth information about those questions which caused most difficulties for answering (e.g. Most time needed to answer, most corrections of initial answer ect.). But that assumes additional considerations mentioned in connection to the new Law on Electronic Communications in Slovenia excepted in 2013.

Theoretical base is drawn down mainly from foreign sources so the question is whether the assumptions of foreign literature can be directly tested in (for discussed online survey) chosen environment (the study is limited to Croatian organizations). But on the other hand this is also one of the main contribution of this paper due to lack of researcher and paper on topic of data about data in survey research in non-English speaking environment (in our case Croatian).

Although Couper (1998) originally coined the term paradata as a general notion for by-product process data which sticks to data about data (not just paradata) despite development of area, we could argue that in further researches of all types of data about data the set of data called now considered as by-products will decrease. Why? With further much needed research it will be more clear which concrete paradata are useful and appropriate for use in connection with what (with which product data or survey variables). That means it is needed better knowing which and where in terms of content paradata can be useful, so that they will be collected and treated intentionally not just as byproduct. For example, which paradata could be potentially useful in order to examine whether the observations in the case of sensitive issues really reflects respondents actual opinion or is it merely the result of answering on quickly and superficially read question without a clearly articulated views on the subject at which the question asks. With knowing the possible usefulness of concrete paradata the set of 'by-product paradata' will be smaller. There will be only those paradata for which we will not know what their specific usable value and another important group of paradata, the one that will eventually be used to respond to the behavior of the respondent

when answering survey. For example paradata which will be used to encourage the respondent to participate till the end of questionnaire. If we draw a line, we can find necessary to conduct additional research on concrete applications of concrete paradata in connection to survey variables or primarily survey product data. Firmer link of paradata with content analysis is essential issue of testing the degree to which paradata are useful. Additional research is also needed to identify possible usefulness of paradata do activate and address some kind of motivational respond to the behavior of the respondent during participation in survey in light of problematic area of dropouts and decreasing response rates.

References

Casas-Cordero C., F. Kreuter, Y.Wang & S. Babey. (2013). 'Assessing the measurement error properties of interviewer observations of neighborhood characteristics.' *Journal of the Royal Statistical Society* 176 (1): 227–249.

Couper, Mick P. and Frauke Kreuter. (2013). 'Using paradata to explore item level response times in surveys.' *Journal of the Royal Statistical Society* 176 (1): 271–286.

Couper Mick P., Eleanor Singer. 2013. 'Informed Consent for Web Paradata Use.' *Survey Research Methods*, 7(1): 57-67.

Horwitz, Rachel, Jennifer Guarino Tancreto, Mary Frances Zelenak & Mary Davis. (2012). *Use of Paradata to Assess the Quality and Functionality of the American Community Survey Internet Instrument*. U.S. Census Bureau: Washington.

Scheuren, Fritz. *Macro and Micro Paradata for Survey Assessment*. Available at:

<http://www.unece.org/fileadmin/DAM/stats/documents/2000/11/metis/crp.10.e.pdf> . (december 2013).

Frost Hubbard, Ben Duffey. (2010). *Paradata, Metadata, Auxiliary Data and other Data about Data: Studying Survey Costs and Errors in the New Millennium*. Matt Jans U.S. Census Bureau: Michigan.

Heerwegh, Dirk and Geert Loosveldt. (2002). 'An evaluation of the effect of response formats on data quality in Web surveys.' *Social Science Computer Review* 20 (4): 471-484.

Heerwegh, Dirk. (a) *Uses of Client Side Paradata in Web Surveys*. Available at:

<https://perswww.kuleuven.be/~u0034437/public/Files/Heerwegh%20Uses%20of%20Client%20Side%20Paradata%20in%20Web%20Surveys.pdf> (December 2013).

Kdaj lahko uporabimo piškotke? Smernice Informacijskega pooblastenca. 2013. Available at: https://www.ip-rs.si/fileadmin/user_upload/Pdf/smernice/Smernice_o_uporabi_piskotkov.pdf (December 2013).

Kreuter, Frauke, Mick Coupery, Lars Lybergz. (2010). The use of paradata to monitor and manage survey data collection. Section on Survey Research Methods – JSM: Maryland.

Nicolaas, Gerry. (2011). Survey Paradata: A review. National Centre for Research Methods: Southampton.

O'Reilly, Jim. Paradata and Blaise: A Review of Recent Applications and Research. Available at: <http://ibuc2009.blaiseusers.org/papers/7d.pdf> (December 2013).

Pirc Musar, Nataša. (2013). Informacijski pooblaščenec izdal smernice glede uporabe piškotkov. 2013. Assessable on: <https://www.ip-rs.si/novice/detajl/informacijski-pooblastenec-izdal-smernice-glede-uporabe-piskotkov/?cHash=5842ad1118a2ae1915f350cc1aa98c22> (December 2013).

U.S. Department of Health and Human Services. (2012). 2011 National Health Interview Survey (NHIS) Paradata File: Public Use Data Release. Hyattsville: Maryland.

Wagner, James. (2009). Paradata and Alternative Measures of Survey Data Quality. University of Michigan: Michigan.

FIGURES AND TABLES

Table 6 Time needed to answer full questionnaire or each page.

Type of organization	Full questionnaire		1st page		2nd page		3rd page		4th page		5th page		6th page	
	Big	Small	Big	Small	Big	Small	Big	Small	Big	Small	Big	Small	Big	Small
N	62	79	62	79	62	79	62	79	62	79	62	79	62	79
Estimated time (min)	15,00	15,00	2,80	2,80	3,00	3,00	3,00	3,00	1,80	1,80	2,10	2,10	0,61	0,61
Mean time needed	15,127 4	15,746 6	4,6207	4,3608	2,7065	2,8152	3,0696	2,6835	1,8167	1,4620	1,534 7	2,5443	1,3793	1,8808
Std. Error of Mean	,87565	,76659	,42282	,21315	,29536	,14712	,49509	,16522	,21264	,11368	,0895 6	,43613	,15210	,31882
Median	13,408 3	14,583 3	3,5333	3,8000	2,2250	2,5833	2,5167	2,3000	1,1750	1,2500	1,400 0	1,5333	,9083	,9167
Mode	12,93 ^a	16,37	2,65 ^a	2,20 ^a	2,23	2,40 ^a	1,58	1,62	,85	1,48	1,13 ^a	1,32 ^a	,63	,67
Std. Deviation	6,8948 8	6,8136 3	3,3292 9	1,8945 2	2,3256 7	1,3076 6	3,8983 5	1,4685 0	1,6743 0	1,0104 0	,7051 9	3,8764 2	1,1976 4	2,8337 4
Variance	47,539	46,426	11,084	3,589	5,409	1,710	15,197	2,156	2,803	1,021	,497	15,027	1,434	8,030
Skewness	1,923	1,152	2,880	1,262	5,531	1,668	6,808	1,499	2,585	2,718	2,385	4,438	1,704	2,806

Std. Error of Skewness	,304	,271	,304	,271	,304	,271	,304	,271	,304	,271	,304	,271	,304	,271
Kurtosis	4,551	1,379	9,573	1,911	36,746	5,317	50,397	2,091	7,108	9,279	7,847	21,123	2,190	7,013
Std. Error of Kurtosis	,599	,535	,599	,535	,599	,535	,599	,535	,599	,535	,599	,535	,599	,535
Minimum	6,13	5,65	1,32	1,68	,82	,70	1,12	,15	,43	,32	,47	,07	,30	,05
Maximum	42,55	39,13	18,97	11,35	18,62	8,97	31,82	7,73	9,15	6,12	4,45	25,37	5,25	12,70

a. Multiple modes exist. The smallest value is shown