Working Papers
www.mmg.mpg.de/workingpapers

Max Planck Institute for the Study of Religious and Ethnic Diversity

Max-Planck-Institut zur Erforschung multireligiöser und multiethnischer Gesellschaften

MMG Working Paper 12-21 • ISSN 2192-2357

Petermann / Heywood / Hewstone / Hüttermann / Schmid / Schmitt / Schönwälder / Stolle / Vertovec
with TNS Emnid

The "Diversity and Contact" (DIVCON) Survey 2010 (wave 1) - Technical Report

Sören Petermann, Joe Heywood, Miles Hewstone, Jörg Hüttermann, Katharina Schmid,
Thomas Schmitt, Karen Schönwälder, Dietlind Stolle and Steven Vertovec
with TNS Emnid

The "Diversity and Contact" (DIVCON) Survey 2010 (wave 1) TECHNICAL REPORT

MMG Working Paper 12-21

Max-Planck-Institut zur Erforschung multireligiöser und multiethnischer Gesellschaften,
Max Planck Institute for the Study of Religious and Ethnic Diversity
Göttingen

ISSN 2192-2357 (MMG Working Papers Print)

Download: **www.mmg.mpg.de/workingpapers**

MPI zur Erforschung multireligiöser und multiethnischer Gesellschaften
MPI for the Study of Religious and Ethnic Diversity, Göttingen
Hermann-Föge-Weg 11, 37073 Göttingen, Germany
Tel.: +49 (551) 4956 - 0
Fax: +49 (551) 4956 - 170

www.mmg.mpg.de

info@mmg.mpg.de

Abstract

The „Diversity and Contact“ (DivCon) project investigates the consequences of diversity on social interactions between individuals with and without migration background and on selected attitudes. The main empirical component of this project is a survey conducted in neighbourhoods of German cities. This technical report is about the first wave of a longitudinal survey with about 2,500 respondents. The report outlines the sample design of the entire study, the survey implementation, a test of representativeness, and information about themes and operationalisations of the questionnaire.

Authors

SÖREN PETERMANN is Research Fellow at the Max Planck Institute for the Study of Religious and Ethnic Diversity (MMG), Department of Socio-Cultural Diversity, Göttingen.

JOE HEYWOOD is a Masters Student at McGill University.

MILES HEWSTONE FBA is Professor of Social Psychology and Fellow of New College, Oxford University.

JÖRG HÜTTERMANN is Research Fellow at the Max Planck Institute for the Study of Religious and Ethnic Diversity (MMG), Department of Socio-Cultural Diversity, Göttingen.

KATHARINA SCHMID is a Research Fellow and lecturer in Psychology at New College, University of Oxford.

THOMAS SCHMITT was Research Fellow at the Max Planck Institute for the Study of Religious and Ethnic Diversity (MMG), Department of Socio-Cultural Diversity, Göttingen (2010-2012) and has now a lecturer position at the University of Erlangen-Nürnberg, Institute for Geography.

KAREN SCHÖNWÄLDER is Research Group Leader at the Max Planck Institute for the Study of Religious and Ethnic Diversity (MMG), Department of Socio-Cultural Diversity, Göttingen and Extracurricular Professor at the Georg August University Göttingen.

DIETLIND STOLLE is Associate Professor in the Department of Political Science at McGill University and member of the Centre of the Study for Democratic Citizenship in Montreal, Canada.

STEVEN VERTOVEC is Director of the Max Planck Institute for the Study of Religious and Ethnic Diversity (MMG), Göttingen, and Head of its Department for Socio-Cultural Diversity.

Contents

The appendix is available only in the online version:
http://www.mmg.mpg.de/publications/working-papers/2012/wp-12-21/

Appendix I: Codebook

A) Context - survey questions
B) Interactions
C) Outcomes
D) Respondent's background
E) Interview information
F) Survey design
G) Contextual data

Appendix II: Questionnaire in German

Appendix III: Questionnaire in English

1. The project „Diversity and Contact“

Many contemporary European and North American societies have become increasingly diverse. One aspect of such diversity is the plurality of ethnic affiliations, cultural preferences and life experiences linked with immigration. How does the socio-demographic and cultural diversity of societies affect the social interactions of individuals within them? To what extent do ethnicity and national origin constitute boundaries that restrict social interactions? And under what conditions are such boundaries irrelevant or overcome? Politically it is also a burning question whether the existence or non-existence of inter-group interaction is relevant for life chances and the overall integration of societies.

This project is about the consequences of diversity on social interactions between individuals with and without migration background and selected attitudes and trust. We seek to investigate to what extent individuals of native or immigrant background interact with each other, to what extent they do so on different levels of social interaction (more superficial encounters, acquaintance/weak ties, friendship/strong ties), how this is influenced by the residential context, and what consequences on some attitudes this may have.

The study is conducted in German cities. This technical report is about the first wave of a longitudinal survey with about 2,500 respondents. The survey is conducted by TNS Emnid (Bielefeld) and is supplemented by systematic observations and a data base on contextual statistics of the 50 neighbourhoods as well as targeted ethnographies and in-depth interviews in 5 selected neighbourhoods. The methodology of the latter is not reported in this publication, but all publications relating to the project can be found on the Max-Planck-Institute website.

The project is based on the assumption that the residential environment (or neighbourhood) is one context in which individuals experience diversity and that this influences residents' attitudes. The mechanisms through which this happens may include more passive observation and experience or socialization. Furthermore, and following on from contact theory, we assume that direct interaction, or contact, has a particular impact on people's attitudes towards those they encounter and also towards diverse others and diversity more generally. Contact is thus seen as a central mediator between existing diversity and its effects on the individuals. Our panel data will enable us to make progress regarding the causal links: do positive attitudes lead to more contact or more contact to more positive attitudes?

Researchers with academic backgrounds in anthropology, human geography, political science, social psychology and sociology belong to the project team: Steven Vertovec, Karen Schönwälder, Jörg Hüttermann, Sören Petermann (all MPI MMG Göttingen), Thomas Schmitt (formerly MPI MMG, now Erlangen University), Miles Hewstone, Katharina Schmid (both University of Oxford) and Dietlind Stolle (McGill University, Montréal). Joe Heywood has helped compile this technical report.

2. Sample design and sampling procedure

The DivCon-study used a stratified sampling design. The sampling procedure was informed by the following main aims: We wanted to conduct a study in a random sample of urban areas. As no Germany-wide data base of urban areas and their characteristics is available, we needed to select cities first. For the selected cities, we could then create a data base of urban areas with their key characteristics from which we could select our areas of investigation.

We wanted to ensure that cities of different size were included in our sample. The experience of diversity might differ between cities of e.g. 50,000 and 500,000 inhabitants. A non-stratified sample drawn from all German cities would have led to an almost exclusive sample of the more numerous small towns. To avoid that we stratified according to three municipality classes: medium-sized towns of 50 to 99,999 inhabitants, big cities of 100 to 499,999 inhabitants and metropolitan cities of 500,000 and more inhabitants.

In order to be able to systematically compare the effects of varying compositions of the population, areas were stratified by share of foreigners. Share of foreigners is the only generally available indicator of immigrant share for small areas. For the respondents in our survey, we collected more detailed information e.g. on former citizenship and parents' country of birth. For populations of urban areas such information is only sometimes available and not in a standardized form across the country.

In debates about trust and "social cohesion" it is often argued that outcomes are mistakenly related to ethnic diversity while in fact they are attributable to deprivation (e.g. Laurence 2011).We used purposive sampling to disentangle effects of socioeconomic conditions and of immigration-related diversity. Here we stratify on share of foreigners and unemployment ratio separately, i.e. overrepresent areas with unusual combinations (e.g. high share of foreigners and low unemployment ratio).

Further, for a representation of the immigrant population in our sample, we mainly relied on the random sampling of respondents. We are not specifically interested in immigrants or particular immigrant groups but in society as a whole, thus we did not have to draw separate samples. However, the selection of many high diversity areas ensures that people with immigrant background are well-represented in the survey.

The survey was conducted by telephone. Because no sampling company offered personal interviews at anything near a realistic price, this choice was not available.

2.1 Description of the population

The population under study is defined by residence and age regardless of citizenship or language. It covers people residing in West German towns and cities of at least 50,000 inhabitants who are of adult age (18 years or older). We restricted our study population to West Germany because of the recent history and low level of ethnic diversity in East Germany, which would have limited our analysis of inter-ethnic interactions and contextual effects of diversity. We restricted our study population to towns and cities of at least 50,000 inhabitants due to limited availability of contextual data for many small towns and rural municipalities. Our study population reflects the majority of the West German urban population and close to half of West Germany's adult population.

We used official statistics as per 31 December 2008 to describe our study population and to draw our survey sample *(Statistisches Bundesamt* 2009). Our study population comprises 24,613,240 adults distributed over 165 municipalities (table 1). About two fifths reside in cities with 500,000 or more inhabitants (called metropolitan cities), which represent 7% of the municipalities. Another two fifths live in

Table 1: Population and municipalities per municipality classification (31/12/2008)

municipality classification	18+ population		municipalities	
	number	%	number	%
metropolitan cities (500,000 and more)	10,295,168	41.8	12	7.3
big cities (100,000 to 499,999)	9,015,074	36.6	58	35.1
medium-sized towns (50,000 to 99,999)	5,302,998	21.6	95	57.6
total	24,613,240	100.0	165	100.0

cities between 100,000 and 499,999 inhabitants (big cities), a category that accounts for over 35% of the municipalities. The final fifth lives in cities of 50,000 to 99,999 inhabitants (medium-sized towns), the majority of all municipalities.

In addition to the distribution of the general population under study, the share of foreigners is also of interest, since this serves as a proxy measure of ethnic diversity. Table 2 shows the total share of foreigners among the whole population (grand mean) and the mean average share of foreigners across municipalities (city mean). Both figures are broken down by municipality category in this table. Of our study population, 14% have a non-German nationality. But foreigners are not equally distributed over the three categories of municipality. The share of 18+ foreigners is above the average in metropolitan cities (17%) and below-average in medium-sized towns (11%). For city means, municipalities with fewer inhabitants are given more weight. Table 2 indicates two tendencies. For metropolitan cities, diversity decreases with population size. For big cities, diversity increases with population size

The average of the city mean is 2% lower than the grand mean because there are much more medium-sized towns than bigger cities. For both values, the share of foreigners is not equally distributed over municipalities of different size - metropolitan cities have the highest share (17%) and medium-sized towns the lowest (11%). The share of foreigners is roughly the same for the grand mean and city mean in medium-sized towns. But the city mean is slightly lower for big cities and slightly larger than the grand mean for metropolitan cities. This is because big cities with fewer inhabitants tend to have lower shares of foreigners, while metropolitan cities with fewer inhabitants tend to have higher proportions of foreigners residing in them.

Table 2: Foreigners per municipality classification (31/12/2008)

	18+ foreigners	share of 18+ foreigners	
		grand mean	city mean
metropolitan cities (500,000 and more)	1,707,336	16.6%	17.0%
big cities (100,000 to 499,999)	1,230,853	13.7%	13.3%
medium-sized towns (50,000 to 99,999)	578,123	10.9%	10.8%
total	3,516,312	14.3%	12.2%

2.2 Sampling stages

It is crucial to our sample design that people are nested within residential areas (*Wohnviertel)* as we aimed to investigate the effects of contextual diversity on individual interactional behaviour and societal attitudes. The sample design includes 50 respondents in each of 50 *Wohnviertel* resulting in a total of 2,500 respondents. To arrive at the ultimate set of cities and residential areas for the survey, we undertook a four-stage random sampling procedure which selected, in turn: cities, residential areas, telephone households and individuals.

First stage: a random selection of 16 cities

a) The sample design is set to 50 respondents in each of the 50 areas (*Wohnviertel)*. With an equal probability selection method based on the adult urban population, the 50 *Wohnviertel* in our study would have comprised 11 from medium-sized towns (because 21.6% of the adult urban population lives in medium-sized towns, see table 1), 18 from big cities (36.6%) and 21 from metropolitan cities (41.8%).

b) The number of *Wohnviertel* per city should correspond to the city's population size, i.e. if the number of inhabitants of city A is twice as high as the number of inhabitants of city B then the number of *Wohnviertel* of city A should be double that of city B. Hence, the number of cities was deduced from the average adult population sizes of the municipality classes. Based on the 18+ population and the number of localities per municipality class (table 1), the average population sizes are 55,821 for medium-sized towns, 155,432 for big cities, and 857,931 for metropolitan cities. Assuming that the sizes of *Wohnviertel* do not vary much between all cities, *Wohnviertel* should be selected according to a 1 to 3 to 15 ratio for medium-sized, big and metropolitan cities respectively.

A good approximation between a) the equal probability sample of 11-18-21 residential areas and b) the 1-3-15 ratio results in 16 cities that comprise 8 medium-sized towns (8 areas), 6 big cities (18 areas) and 2 metropolitan cities (24 areas).

In order to ensure a representative sample of cities as primary sampling units, the sample frame was stratified by municipality class, share of foreigners and region. Firstly, we used three strata by municipality class: medium-sized towns between 50,000 to 99,999 inhabitants, big cities between 100,000 and 499,999 inhabitants, and metropolitan cities with 500,000 or more inhabitants. Secondly, in order to ensure sufficient variance of contextual ethnic diversity, municipalities were sorted in descending order of share of foreigners within each municipality class. The resulting lists of

Table 3: Sampled cities (31/12/2008)

city	18+ population	18+ foreigners in %	region
metropolitan cities (500,000 and more)			
Frankfurt am Main	563,113	20.8	south
Hamburg	1,500,346	14.2	north
big cities (100,000 to 499,999)			
Mannheim	263,431	23.6	south
Ingolstadt	102,376	15,9	south
Krefeld	196,874	13.1	west
Bochum	323,022	11,2	west
Leverkusen	133,764	10.8	west
Lübeck	177,598	8.1	north
medium-sized towns (50,000 to 99,999)			
Schweinfurt	45,259	15.2	south
Konstanz	71,192	13.9	south
Gießen	64,310	11.6	south
Herten	52,551	10.3	west
Dormagen	51,949	9.9	west
Delmenhorst	61,525	8.6	north
Viersen	62,353	7.8	west
Emden	42,537	7.1	north

municipalities were divided into groups. The number of groups differs between the three municipality classes according to the number of cities that has to be sampled, i.e. eight groups of medium-sized towns, six groups of big cities, and two groups of metropolitan cities. The third stratification criterion was region. The three regional strata are the north with Schleswig-Holstein, Hamburg, Bremen, Lower Saxony, Berlin, north-eastern North Rhine-Westphalia and Kassel in Hessen, the west with south-western North Rhine-Westphalia and northern Rhineland-Palatinate, and the south with southern Rhineland-Palatinate, Hessen without Kassel, Saarland, Baden-Württemberg, and Bavaria. The stratification scheme has 16 municipality class * share of foreigner strata. The regional stratification triples the stratification scheme.

We selected a region and then a city. The selection of regional cells depended on the regional distribution of the adult population per municipality class and was weighted by its adult population. Once the 16 stratification cells were selected, municipalities within the cells were then selected by random sampling using a research randomizer (**www.randomizer.org/form.htm**). In the selected cells, a city was sampled randomly weighted by its population size. The sampling procedure took two steps. Firstly, all elements were arranged in a random order. Secondly, elements were randomly drawn. 1 independent set was drawn. See table 3 for the results of random city sampling.

Second stage: a random selection of 50 residential areas

At the second stage, a random sample of 50 *Wohnviertel* was drawn. These areas are sub-city units according to official definitions by the respective municipality. The designation of the areas differs in the 16 municipalities between *Stadtteile, Stadtbezirke, statistische Bezirke* and *Sozialräume*. A *Wohnviertel* has about 7,000 inhabitants, on average. If the population size was below the minimum of 2,000 inhabitants, we either merged it with a neighbouring area or excluded the area from the sample frame. If the population size exceeded the maximum of 14,000 inhabitants, we partitioned the area into sub-units if the required statistics were available for those sub-units. Most of the cities were treated with that kind of area adaptation, except Delmenhorst, Herten, Ingolstadt and Leverkusen.

The *Wohnviertel* of each city are pools of areas from which we randomly drew our 50 *Wohnviertel*. The biggest pool is Hamburg with 190 areas and the smallest pools are Gießen and Emden with only 8 areas. According to the above-mentioned rule, we selected 1 *Wohnviertel* per medium-sized town (=8 *Wohnviertel*), 3 *Wohnviertel* per big city (=16 *Wohnviertel*) and 12 *Wohnviertel* per metropolitan city (=24 *Wohnviertel*).

In order to increase the variance of ethnic diversity across our selected *Wohnviertel*, we employed purposive sampling based on two dimensions, meaning that we used ethnic diversity and socio-economic characteristics to stratify *Wohnviertel*. Ethnic diversity is represented by the share of foreigners, and socio-economic status is represented by the unemployment rate.[1] Both measures were dichotomized on the respective median value of every city. That gives half of the *Wohnviertel* with low ethnic diversity and the other half with high ethnic minority and half of the *Wohnviertel* with low socio-economic status and half with high socio-economic status (where a high unemployment rate means a low socio-economic status and vice versa).

1 Statistics on socio-economic status for *Wohnviertel* are rarely available. Statistics on income tax exist only for the municipality level.

Table 4: Overview of *Wohnviertel* in the sampled cities

city	number	designation	average size
metropolitan cities (500,000 and more)			
Frankfurt am Main	103	Stadtbezirke	6,503
Hamburg	190	Stadtteile	9,048
big cities (100,000 to 499,999)			
Mannheim	40	Statistische Bezirke	7,333
Ingolstadt	12	Statistische Bezirke	10,322
Krefeld	41	Statistische Bezirke	5,837
Bochum	42	Statistische Bezirke	8,813
Leverkusen	16	Statistische Bezirke	10,116
Lübeck	27	Stadtbezirke	7,626
medium-sized towns (50,000 to 99,999)			
Schweinfurt	13	Stadtteile	4,071
Konstanz	11	Stadtteile	6,826
Gießen	8	Stadtteile	8,632
Herten	9	Stadtteile	7,067
Dormagen	12	Stadtteile	5,242
Delmenhorst	10	Stadtteile	7,578
Viersen	16	Sozialräume	4,554
Emden	8	Stadtteile	6,278

Table 5: Stratification scheme of residential areas

Municipality class		Medium-sized towns (1 area)		Big cities (3 areas)		Metropolitan cities (12 areas)	
Ethnic diversity		low	high	low	high	low	high
Socio-economic status	low	¼	¼	½	1	2	4
	high	¼	¼	½	1	2	4

Combined with the three existing strata of municipality class this resulted in a 3x2x2 stratification scheme for the selection of *Wohnviertel* (table 5). For each selected medium-sized town, big city and metropolitan city, we selected one, three and twelve areas respectively (proportional stratification). Additionally, high diversity areas in big and metropolitan cities were oversampled. Thus, two out of the three *Wohnviertel* per big city and eight out of the twelve *Wohnviertel* per metropolitan city are high ethnic diversity areas.

While the stratification criteria of high/low ethnic diversity and high/low socio-economic status are the same for all three kinds of municipalities, the selection process differed. There was a two step selection process for medium-sized towns. Firstly, there was a random selection of one stratum, where each combination of ethnic diversity and socio-economic background was selected twice in the 8 medium-sized towns. Secondly, there was a random selection of one *Wohnviertel* of the selected stratum for each medium-sized town. The same selection logic was applied for low ethnic diversity *Wohnviertel* in big cities. First, there was a random selection of one stratum, where each socio-economic status was selected three times in the 6 big cities. Secondly, there was a random selection of one *Wohnviertel* of the selected stratum, one of the high ethnic diversity/low socio-economic status and one of the high ethnic diversity/high socio-economic status in the 6 big cities. For metropolitan cities, two or four *Wohnviertel* are selected for the four strata (as indicated in table 5). The resulting selection of 50 areas includes 18 low and 32 high ethnic diversity areas as well as 24 low and 26 high socio-economic status areas.[2]

Table 6: Overview of the 50 selected *Wohnviertel*

characteristic	min	max	median	average	standard deviation
population	2,826.0	18,611.0	7,155.0	7,571.0	3,205.49
share of foreigners	1.9	46.3	15.0	16.3	9.65
unemployment ratio	1.2	10.2	5.0	5.0	1.96

Table 6 gives an overview of the 50 selected *Wohnviertel*. The population of the 50 Wohnviertel is on average approx. 7,500 people, with a wide range from 2,800 to

2 In one of the cities, there was no low ethnic diversity/low socio-economic status area. We switched to a low ethnic diversity/high socio-economic status area instead.

13,000 people. One outlier has more than 18,500 inhabitants. Share of foreigners is 16% on average with a wide range from 2% to 46%.

Third and fourth stage: a random selection of private landlines and respondents

The third stage was about selecting private telephone households within the *Wohnviertel*. This is especially tricky because not all people can be reached by telephone, some people are only reachable by cell phones and often telephone numbers are not listed in telephone books. We discarded the problem of non reachability by landlines because only a small percentage of people do not have a landline. In 2007, 92% of the German population could be reached by a landline and 7% only by mobile phones (**www.bik-gmbh.de**). Less than 2% had no telephone connection. Non-listed landlines are a much more troublesome problem. We therefore decided to draw 60% of the sample from the telephone register and 40% from generated telephone numbers using the random digit dialing technique (RDD) based on the telephone register. The two subsamples were checked for double numbers.

The fourth and final stage was the selection of one respondent within a private telephone household. We applied a Kish grid for a strictly random selection, where the contact person is asked about the number of household members aged 18 or older. A random procedure then selects the oldest/second oldest/third oldest person in the household as respondent. This selection is done by a computer. The interviewer has no influence on it.

3. Development of the questionnaire

3.1 Questionnaire

The questionnaire for the 2010 DivCon-survey was developed by the research team, tested and revised for the main survey. Several related questionnaires were analysed for common and tested conceptualisations of particular issues. This included those of the US study "Citizenship, Involvement, Democracy", the "European Social Survey", the "International Social Survey Programme", the Canadian study "Connected Lives" 2004, the "Social Capital Community Benchmark Survey", the "Ethno-religious Diversity and Social Trust in Residential and Educational Settings" project, and the "*Allgemeine Bevölkerungsumfrage der Sozialwissenschaften"* (Allbus).

3.2 Cognitive interviews

In order to find the most suitable wording for three aspects of our questionnaire, cognitive interviews were conducted. Two group conversations took place on 9 and 11 March 2010 in Göttingen. The participants were fairly mixed in terms of age, gender and migration background. The interviews lasted about 75 minutes.

First, we tested how best to refer to the residential area. Terms like „*Nachbarschaft*", „*Wohnviertel*", „*Wohngegend*", „*Stadtteil", „Ortsteil*" were discussed. As for instance "*Nachbarschaft*/neighbourhood" tends to refer to the immediate environment of neighbouring houses and flats, "*Wohnviertel*/*residential quarter*" came out as the most appropriate term.

Second, we tested how best to ask about personal networks. Here we were not only interested in the terminology but also in the extent to which people can give numbers of friends and acquaintances. Consequently, we used approximations (e.g. "between 10 and 20") for the circle of acquaintances in the survey.

Third, we tested how best to refer to the immigrant and non-immigrant population. Here, terms like „*Migranten*/(im-)migrants", „*Ausländer*/foreigner", „*Personen, die selbst und deren Eltern nicht aus Deutschland stammen*/people who are not themselves native Germans or whose parents are not from Germany" for the immigrant population and terms like „*Nicht-Migranten*/non-migrants", „*alteingesessene Deutsche*/native Germans", for the non-immigrant population were discussed.

3.3 Questionnaire pilot test

A draft questionnaire developed by the project team was pretested between 29 March and 10 April 2010 in the Emnitel-telephone studio in Göttingen. Team members were present. Altogether 79 interviews were conducted with residents of Bonn and Ludwigsburg, i.e. cities not included in the survey (one residential area in each city).

One aim of the pretest was to find out to what extent the use of computer generated numbers was feasible. It is not possible in Germany to deduce from telephone numbers whether people live in particular areas of a city. Thus a huge number of people have to be called and asked where they live in order to find respondents in the selected areas. As only five per cent of the computer generated numbers turned out to lead to respondents in designated areas, we decided to also use registered numbers from the telephone book where usually a street address is given.

On the basis of the pretest, the questionnaire was revised.

3.4 Questionnaire translations

We aimed to include as many immigrant respondents as possible in the survey. Hence, the questionnaire was translated into six languages that cover the largest immigrant groups in German cities: Turkish, Russian, Polish, Italian, Serbo-Croat and English. Translations were retranslated into German to check for correctness and completeness. Translations in both directions were done by professionals.

4. Survey implementation (with TNS Emnid)

4.1 Interviewer

The 338 CATI-Interviewers included some who could offer to conduct the interview in one of the six non-German languages (Turkish, Russian, Polish, Italian, Serbo-Croat, English).

The interviewers took part in a training specifically for this project before the beginning of the field phase and training continued during throughout the field phase. One focus was how to enquire about the street name and the location of the household in a survey area. Other aspects of the training concerned the introduction of the institution conducting the survey, the aims of the study and specific questions. Interviewers were also supplied with written material they could consult if necessary.

In the telephone studio supervisors were continuously present. They were familiar with the study and could intervene if problems or questions came up, but also supervised the correct conduct of the interviews. On average one supervisor was assigned 15 interviewers. Because of computer-aided interviewing, supervisors could permanently control the interviewing process. They are able to listen in and to monitor how the interviewers fill in their forms.

4.2 Cover letter

Households selected from telephone books received a letter announcing the survey. In order to increase the willingness to take part in the survey, the letter explained the aims and intentions of the study in a general and accessible way. Letters also explained data protection issues and ensured the recipients that no personal information would be passed on to others not involved in the study. Immediately before the start of the interviewing process, on 19 May 2010, 6.200 letters were posted. A month later, on 15 June 2010, almost 2.800 letters were mailed. Later on, further addresses were drawn and letters sent accordingly.

4.3 Fieldwork

The field phase lasted from 17 May to 20 July 2010. Interviews on average lasted 40.3 minutes. As common for complex studies, the timespan varied considerably, from 17 to more than 120 minutes. It is unknown, however, whether interruptions

were the cause of longer interviews. Ten per cent of the interviews were completed after 31 minutes, 90 per cent after 51 minutes. Only five interviews took more than 90 minutes.

4.4 Response rate

The overall response rate for this study was calculated to be 24.3 per cent (table 7). This is not a-typical for telephone surveys. Further we have to take into account that the need to ask relatively detailed screening questions (to ensure that respondents lived in the target areas) discouraged some potential respondents.

Of the selected and generated telephone numbers, several turned out to be invalid. This was the case for 60 per cent of the generated numbers and around 15 per cent of the numbers from the telephone book. 13,567 generated numbers and 831 phone book numbers were not used because the target of 50 interviews had been achieved. Thus the actually used adjusted gross sample consisted of 215,495 generated numbers and 9.769 phone book numbers.

A significant share turned out to be 'out-of-scope', i.e. households did not live in the selected areas (72%). Other neutral losses were due to inability to make contact (after 15 attempts), the number belonging to a business, inability to conduct the interview in one of the six languages offered (854 cases), and illness. Without those neutral losses, of the generated numbers, only 2 per cent remained, and 56% of the telephone book numbers. Thus the net sample comprised 4824 generated and 5507 phone book numbers.

Refusals and incomplete interviews add up to 7825 cases (3821 generated numbers + 4004 from the phone book). Table 7 distinguishes refusals at the household level and at the level of the known-respondent. Based on the net sample, 20.8 per cent of the generated numbers and and 27.3 per cent of the phone book numbers led to successful interviews. This may be seen as a rather low response rate, but we have to take into account that a complex screener was set up to check whether the household belongs to the target area and to identify the person to be interviewed within the household. Additionally, cover letters that decrease refusals at the household level, could only be sent to households of the phone book sample.

If we compare the systematic losses for the generated and the phone book numbers, we can see that the share of refusals at the level of the known-respondent is nearly the same for both categories. The same is true for incompleted interviews.

Table 7: Survey response

	total phone numbers		generated numbers		phone book numbers	
	absolute	per cent	absolute	per cent	absolute	per cent
total phone numbers	588,048		575,590		12,458	
not assigned phone numbers	348,386		346,528		1,858	
gross sample	239,662		229,062		10,600	
not attempted	14,398		13,567		831	
total phone numbers used (adjusted gross sample)	225,264	100%	215,495	100%	9,769	100%
neutral losses	214,933	95.4%	210,671	97.8%	4,262	43.6%
no private household	8,566	3.8%	8,121	3.8%	445	4.6%
nobody in target group	204	0.1%	189	0.1%	15	0.2%
answering machine/ free line/always busy	17,597	7.8%	16,176	7.5%	1,421	14.6%
fax/modem	19,037	8.5%	18,752	8.7%	285	2.9%
respondent unavailable during field period	4,952	2.2%	4,126	1.9%	826	8.5%
respondent illness	1,483	0.7%	1,020	0.5%	463	4.7%
respondent language problem	854	0.4%	740	0.3%	114	1.2%
out of sample area	162,240	72.0%	161,547	75.0%	693	7.1%
net sample	10,331	4.6%	4,824	2.2%	5,507	56.4%
net sample	10,331	100%	4,824	100%	5,507	100%
systematic losses	7,825	75.7%	3,821	79.2%	4,004	72.7%
household refusal	5,873	56.9%	2,905	60.2%	2,968	53.9%
respondent refusal	1,707	16.5%	806	16.7%	901	16.4%
break off	245	2.4%	110	2.3%	135	2.5%
complete interviews	2,506	24.3%	1,003	20.8%	1,503	27.3%

Refusals at the household level are less common for the phone book numbers. This may be attributable to the introductory letters sent to these households before the telephone contact.

Of the 245 interviews that could not be completed a quarter were ended in the first three minutes, when interviewers asked questions about the street address (to confirm the location of the household in the survey area) and aimed to select the interview partner. Otherwise, there is little identication that particular questions led to unsuccessful interviews.

51 (2%) of the 2506 complete interviews were conducted at least partially in one of the six foreign languages. 25 of these interviews were realized in Russian, 9 in Turkish, 8 in Polish, 5 in Serbo-Croat, 3 in Italian and 1 in English.

5. The composition of the sample population – representativity and weighting

As a result of the multi-stage sampling procedure and unit non-response, the sample population might be biased. Unequal probabilities of selection due to clustering and individuals refusing to participate might cause discrepancies between the sample and the study populations. First, we provide a test of representativity for the DivCon 2010 survey data to control for these discrepancies. Second, we describe the sampling weights included in the data set to correct for sample bias.

The test of representativeness of the DivCon 2010 survey data comes in the form of a comparison with the 2008 *Mikrozensus*. This sample census covers 1% of the total population in Germany (roughly 800,000 people). While it is a survey itself, there is an obligation on the part of participants to respond to it. This census is carried out by the *Statistische Landesämter* and the *Statistisches Bundesamt* (Federal Statistical Office). It is weighted in key socio-demographic variables, so does not deviate significantly from the population as a whole.

In Tables 8, 9 and 10 we show how our sample population compares with the *Mikrozensus* population in terms of gender and age, indicators of migration background, education and income. In the "difference" column, the percentage that the value for the sample population differs from the total study population represented by the *Mikrozensus* is set out. Frequency distributions and differences are listed for all respondents and separated into three municipality classes of medium-sized

towns (50,000 to 99,999 inhabitants), big cities (100,000 to 499,999 inhabitants) and metropolitan cities (500,000 and more inhabitants). When the difference between the sample population and the whole population is greater than five percent, we regard this as problematic. In tables 8, 9 and 10, such differences are highlighted.

Socio-demographic groups resulting from combinations of gender, age and nationality are presented in table 8. The categories most underrepresented in the sample are German men aged 20-44, especially in larger cities. This age group is underrepresented by over 4% in medium-sized towns and big cities and by over 5% in metropolitan cities. German women older than 45 years are the most over-represented in the sample.

Table 9 shows figures for migration background. As distinct from other surveys, with 21% of the respondents, people with migration background are not strongly underrepresented despite the fact that we had only an indirect sampling procedure to boost this group. However, the foreign-born and foreigners among the people with migration background are underrepresented.

Table 10 shows data for socio-economic status. Those who finished school having completed 12 or 13 years of education with a higher school certificate are strongly overrepresented by over 13% compared to the microcensus data. Participants who completed secondary school after 8 years are heavily underrepresented in the survey. A similar picture emerges for income with those earning 900/1000 to 1500 Euros per month being underrepresented and those making 2000 to 2900/3000 Euros being overrepresented by just over 5%. This figure is most pronounced in metropolitan cities. Compared to the microcensus, participants refusing to report their income are also overrepresented. But this might be due to voluntary (DivCon) and obligatory (Microcensus) income statements.

The sample weights, constructed by the Max Planck Institute Research Team, adjust for key variables of interest to make the sample population conform to the study population. Posterior weights were computed on the basis of nominal-actual comparisons. We use two sources for these comparisons: the 2008 microcensus for the total sample population and 2009 area statistics for each subsample population of the 50 *Wohnviertel*. The key variables identified for this report are age, gender, nationality, education and migration background. The purpose of the weights included in the data set therefore is to make the sample population equivalent on these variables to allow for the estimation of population characteristics and sampling errors.

Six weights are based on the 2008 microcensus according to municipality classification (medium-sized towns, big cities, metropolitan cities), gender (female, male),

Table 8: Socio-demographic background

			all municipalities		municipalities with at least … inhabitants					
					50,000 to 99,999		100,000 to 499,999		500,000 and more	
gender	nationality	age	per cent	difference	per cent	difference	per cent	difference	per cent	difference
male	Germans	18 to 24 years	1.9%	-2.5%	2.0%	-2.6%	2.0%	-2.5%	1.7%	-2.3%
		25 to 44 years	10.3%	-4.3%	12.3%	-1.5%	10.3%	-4.1%	9.6%	-5.5%
		45 to 64 years	15.5%	2.6%	16.5%	2.4%	16.0%	3.1%	14.9%	2.7%
		65 years and more	12.7%	3.6%	11.3%	1.0%	12.8%	3.6%	13.1%	4.7%
	foreigners	18 to 24 years	0.1%	-0.7%	0.0%	-0.7%	0.1%	-0.8%	0.2%	-0.7%
		25 to 44 years	0.7%	-3.0%	0.8%	-2.0%	0.8%	-2.6%	0.6%	-3.8%
		45 to 64 years	0.7%	-1.4%	0.5%	-1.1%	0.3%	-1.6%	1.1%	-1.4%
		65 years and more	0.2%	-0.6%	0.3%	-0.3%	0.3%	-0.4%	0.1%	-0.8%
female	Germans	18 to 24 years	2.2%	-2.2%	3.0%	-1.6%	2.5%	-2.3%	1.8%	-2.3%
		25 to 44 years	11.9%	-2.3%	13.0%	-0.5%	11.3%	-2.7%	11.9%	-2.8%
		45 to 64 years	21.3%	7.8%	23.8%	8.9%	20.1%	6.6%	21.4%	8.7%
		65 years and more	19.3%	6.7%	15.3%	1.8%	20.7%	7.7%	19.7%	7.7%
	foreigners	18 to 24 years	0.1%	-0.7%	0.0%	-0.7%	0.2%	-0.6%	0.1%	-0.9%
		25 to 44 years	1.4%	-2.3%	1.3%	-1.4%	1.0%	-2.5%	1.7%	-2.6%
		45 to 64 years	1.2%	-0.9%	0.0%	-1.7%	1.1%	-0.8%	1.6%	-0.9%
		65 years and more	0.5%	-0.1%	0.3%	-0.2%	0.3%	-0.2%	0.7%	0.0%
n			2,485		400		890		1,195	

Table 9: Migration background

	all municipalities		municipalities with at least … inhabitants					
			50,000 to 99,999		100,000 to 499,999		500,000 and more	
migration background	per cent	difference	per cent	difference	per cent	difference	per cent	difference
people without migration background	78.9%	4.1%	81.8%	4.6%	78.0%	3.6%	78.5%	4.6%
people with migration background	21.1%	-4.1%	18.2%	-4.6%	22.0%	-3.6%	21.5%	-4.6%
n	2,506		401		900		1,205	

Table 10: Socio-economic background

	all municipalities		municipalities with at least … inhabitants					
			50,000 to 99,999		100,000 to 499,999		500,000 and more	
highest level of school education	per cent	difference	per cent	difference	per cent	difference	per cent	difference
still in school	0.2%	-1.0%	0.5%	-0.7%	0.2%	-0.9%	0.1%	-1.1%
no graduation	1.0%	-4.3%	0.7%	-4.0%	1.3%	-4.3%	0.8%	-4.5%
secondary school certificate (8 classes): Volks-/Hauptschulabschluss	22.9%	-13.2%	23.7%	-19.3%	27.3%	-11.7%	19.4%	-10.6%
secondary school certificate (10 classes): Abschluss der POS; Realschulabschluss	29.0%	6.0%	31.2%	7.0%	29.0%	8.1%	28.3%	4.0%
advanced technical certificate: Fachhochschulreife	9.0%	2.4%	9.2%	2.3%	10.3%	3.4%	8.0%	1.7%
higher school certificate (12/13 classes): Allgemeine/fachgebundene Hochschulreife	37.5%	10.8%	34.7%	15.6%	31.6%	6.1%	43.0%	11.1%
no answer	0.3%	-0.6%	0.0%	-0.9%	0.2%	-0.6%	0.4%	-0.6%
n	2,506		401		900		1,205	

monthly net income (household)	per cent	difference	per cent	difference	per cent	difference	per cent	difference
less than 500 Euro	1.5%	-0.2%	1.0%	-0.5%	1.8%	-0.3%	1.4%	-0.1%
500 to 899 Euro (DivCon: … to 999 Euro)	8.3%	0.2%	7.7%	1.6%	9.7%	1.5%	7.5%	-1.5%
900 to 1.499 Euro (DivCon: 1.000 to …)	12.9%	-5.5%	12.5%	-4.0%	13.8%	-4.2%	12.4%	-7.3%
1.500 to 1.999 Euro	12.8%	-2.3%	11.2%	-3.1%	12.8%	-1.9%	13.3%	-2.5%
2.000 to 2.899 Euro (DivCon: … to 2.999 Euro)	26.9%	5.3%	25.9%	3.5%	25.6%	4.2%	28.1%	6.9%
2.900 to 3.999 Euro (DivCon: 3.000 to …)	11.1%	-3.0%	12.2%	-3.2%	10.1%	-3.6%	11.5%	-2.4%
4.000 to 4.999 Euro	4.5%	-1.5%	4.2%	-2.1%	4.9%	-0.9%	4.3%	-1.7%
5.000 to 7.499 Euro	3.4%	-0.9%	6.0%	1.8%	3.1%	-0.9%	2.8%	-1.9%
7.500 Euro and more	1.6%	0.0%	1.5%	0.2%	1.7%	0.3%	1.5%	-0.4%
no answer	17.1%	7.9%	17.7%	5.8%	16.7%	5.6%	17.3%	11.0%
n	2,506		401		900		1,205	

age groups (18-24, 25-44, 45-64, 65+), nationality (German, foreign), migration background (with, without migration background) and highest level of school education (still in school, no graduation, secondary school certificate, 8 classes: *Volks-/ Hauptschulabschluss*, secondary school certificate 10 classes: Abschluss *der POS/ Realschulabschluss*, advanced technical certificate: *Fachhochschulreife*, higher school certificate 12/13 classes: *Allgemeine/fachgebundene Hochschulreife*, no answer). A seventh microcensus weight combines all these single variable weights. Three weights are based on 2009 area statistics of gender (female, male), age groups (18-24, 25-44, 45-64, 65+) and nationality (German, foreign). One more additional weight combines all three single weights.

For each of these posterior weights, a value of one indicates that the sample and study population are equal. Values of less than one indicate overrepresented cases meaning that these cases have to lose weight compared to the study population. Vice versa, values greater than one give underrepresented cases more weight. However, weights should be carefully used if they are much smaller or much greater than one which is the case for combined DivCon weights in particular. Furthermore, Winship and Radbill (1994) recommend the use of weights if they are a function of the dependent variable. For analyses where weights are a function of the independent variables - as in most DivCon analyses - weighting procedures should be abandoned.

Additional to these post-stratification weights that correct unit non-response, the specific survey design with cities as units of the primary sampling stage and neighbourhoods as the secondary sampling stage needs corrections too. STATA offers a command (`svyset`) to declare cluster and strata identifier as well as population corrections per sampling unit and can include sample weights for the inclusion probability. Once the study design is set, analyses give results for the corrected sample. However, multi-level regressions already take the nested structure of respondents within neighbourhoods into account.

6. Contextual data

Our research design distinguishes a city level, an area level and an individual level. Individuals are nested in 50 areas which in turn are nested in 16 cities. Both locality levels serve as socio-spatial contexts for the individual interactions and attitudinal outcomes. We expect city and area effects of immigration-related diversity and control for other context characteristics. Four different types of contextual information were sought after to meet the requirements of the DivCon projects' aims: immigration-related diversity, population structure, socio-economic conditions and urban structure.

Apart from the interviews, the data set contains data from two sources: The research team conducted area explorations, systematic observations in each of the 50 areas. One or two researchers explored each area for between 3 and 6 hours depending on its size. An area exploration had three elements: a general exploration of the area on foot, an observation of shops and gastronomy, and a systematic count of people at a vivid public place (e.g. a central bus stop). The aim of the area explorations was threefold. Firstly, we used information from all three observational elements to classify an area's noticeable diversity. Secondly, we assessed contact opportunities in public spaces, e.g. shopping zones, playgrounds, parks. Thirdly, we gathered information on the residential building structure. The first variable "noticeable diversity" measures an aspect of immigration-related diversity, while the two other variables are aspects of the urban structure.

Second, city and area level data was collected by the Max Planck Institute from the statistics departments of each one of the sixteen municipalities that the participants in the sample population reside in. In Germany, such data are only partially available from central sources, and often not even collected according to a general standard. The collection was as standardised as possible in terms of gathered/computed statistics and key dates. Several contextual characteristics were found, collected in a database and added as variables to the DivCon data set.

7. Themes and operationalisations in the data set

A) Context - survey questions

A.1) Area context

Individual relevance

Given the focus on interactions between immigrants and non-migrants in the DivCon project, a substantial module in the survey is dedicated to relevance of the neighbourhood for the individual and its perception. First, they were asked, on a five-point scale, the extent to which they feel comfortable in their neighbourhood. They were then asked about the amount of free time (which can include chores like shopping, going to the doctor and so on) they spend in the neighbourhood. The range of options varies from practically all the time, which is coded as one to almost none of the time, which is coded as five. Given that many peoples' interactions can be centred around their work or place of study, a dummy variable is available for whether or not the respondent's place of work, university or school is situated in the neighbourhood they reside in.

Looking at neighbourhood-level data comes with inherent problems of self-selection bias, since a large proportion of residents in a given area decided to move there, rather than being randomly selected. Respondents were asked for the main reason why they decided to move to this particular address in the first place. In the survey, this question is left open, so respondents could answer however they wished. The text responses are available in the data set. But the main categories have also been given a numerical value. They are as follows: job-related reasons, family reasons, contacts in the neighbourhood, accommodation, material conditions and social conditions.

Length of residence

Individuals who have lived in a given neighbourhood for a long time are more likely to have established contacts. Participants in the survey were asked the year that they first moved into the neighbourhood. This has been coded in the data set as the number of years they have lived in the area when the survey was taken. It has also been recoded into four groups (0-20 years, 21-40 years, 41-60 years and 60 or more years). A second recoding gives a less even distribution, but one that is perhaps more intuitive for some studies (0-10 years, 11-20 years, 21-40 years and 41 years or more).

Perceptions of diversity

A number of items in the DivCon survey capture how diverse individuals feel that their neighbourhood is, whether or not they are happy with the perceived level

of diversity and also how friendly they perceive relations in the neighbourhood to be.

Respondents were asked whether they perceive the inhabitants of the neighbourhood they live in to be 'quite diverse' or 'very similar'. This has been recoded to vary from zero for more similar and one for more diverse. Finally, perceptions of relations

A.1) Area context	
	Individual relevance
v3, v3_rec	feel comfortable in the neighbourhood
v4, v4_rec	time spent in the neighbourhood
v7	main reasons for moving to this particular neighbourhood
v7_1_rec	job-related reasons
v7_2_rec	family-related reasons
v7_3_rec	contacts in the neighbourhood
v7_4_rec	accommodation-related reasons
v7_5_rec	infrastructure of the neighbourhood (material conditions)
v7_6_rec	population of the neighbourhood (social conditions)
v40, v40_rec	place of work, school or uni in the neighbourhood
	Length of residence
v6, v6_rec	living in the neighbourhood since the year
residence_n	length of residence in the neighbourhood in years
residence_n_grp	length of residence in neighbourhood in 4 groups
residence_n2_grp	length of residence in neighbourhood in 4 groups
	Perceptions of diversity
v8, v8_rec	perception of general diversity in the neighbourhood
v9	aspects of diversity
v9_1_rec	immigration-related diversity
v9_2_rec	lifestyle-related diversity
v9_3_rec	socio-economic diversity
v9_4_rec	socio-demographic diversity
v9_5_rec	religious diversity
v9_6_rec	diversity related to social behaviour/neighbourliness
v10	relationship between people in the neighbourhood
v10_d1	friendly relations in the neighbourhood
v10_d2	unfriendly relations in the neighbourhood
v10_d3	neutral relations in the neighbourhood
v12, v12_d	perception of immigration-related diversity in the neighbourhood
v13, v13_rec	feeling about immigration-related diversity in the neighbourhood

in the neighbourhood are measured by the question 'How is the relationship between people' with three possible categories: friendly, unfriendly or neither.

In terms of the proportion of natives to immigrants in the neighbourhood, a four point scale is given that ranges from almost exclusively Germans, coded as zero, to mostly Germans coded as one, people from many different countries coded as two and mostly people from other countries coded as three. This is followed up by asking respondents how they feel about this situation on a scale of one for being happy through to five for not feeling good at all about it.

A.2) City context

In addition to the neighbourhood context, some of the questions that were asked about the neighbourhood were also asked about the municipality that the individual resides in, specifically the number of years that an individual has lived there and their perception of diversity in the city. As with the neighbourhood proportion of natives to immigrants variable, the four possible categories are nearly all Germans, mostly Germans, people from many countries and mostly people from other countries.

The municipality class (medium town, large city or metropolitan city - see section 2) is also indicated in the data set.

A.2) City context	
	Length of residence
v5, v5_rec	living in the city since the year
residence_c	length of residence in the city in years
residence_c_grp	length of residence in city in 4 groups
	Perceptions of diversity
v11, v11_d	perception of immigration-related diversity in the city

B) Interactions

B.1) Contact

Contact theory provides the basis for many of the modules included in the DivCon survey. According to contact theory, encounters with out-group members can positively influence perceptions of out-groups.

Inter-group contact was measured with questions on how often respondents talk to immigrants and non-immigrants respectively. This question was asked for two settings: the neighbourhood and the workplace. These items were followed in the survey

by questions that measure quality of contact. Participants were asked how pleasant they found the conversations to have been, on a five point scale varying from one for very unpleasant to five for very pleasant. The survey further asked a general question about contact with members of four groups. These groups include Germans, Turkish, other Western European, and *Russlanddeutsche* (colloquial for ethnic Germans). Being part of one of these groups is specified by either being born in one of those countries, or having parents who were born in that country. Being German means being born in Germany to German parents. The frequency of contact can be chosen from one of five categories: daily, weekly, monthly, less often than monthly, or never.

In addition to this, another question in the survey asks how often (using the same categories as the above variable) the respondent has contact with people who live abroad.

Extent of indirect out-group contact

In addition to direct contact, the *indirect* contact hypothesis posits that knowing that close friends have inter-group friendships can lead to reduced prejudice, since people

B.1) Contact	
	Quantity of direct contact with specific groups
v42_1, v42_1_rec	frequency of contact with Turks
v42_1_rec2	frequency of contact with Turks (0-100 scale)
v42_2, v42_2_rec	frequency of contact with Russlanddeutsche
v42_2_rec2	frequency of contact with Russlanddeutsche (0-100 scale)
v42_3, v42_3_rec	frequency of contact with Western Europeans
v42_3_rec2	frequency of contact with Western Europeans (0-100 scale)
v42_4, v42_4_rec	frequency of contact with native Germans
v42_4_rec2	frequency of contact with native Germans (0-100 scale)
v44, v44_rec	frequency of contact with people living abroad
v44_rec2	frequency of contact with people living abroad (1-100 scale)
	Quantity and quality of direct out-group contact across settings
v45, v45_rec	frequency of contact with out-groupers in the neighbourhood
v45_rec2	frequency of contact with out-groupers in the neighbourhood (0-100 scale)
v46, v46_rec	perception of contact with out-groupers in the neighbourhood
v47, v47_rec	frequency of contact with out-groupers at the workplace
v47_rec2	frequency of contact with out-groupers at the workplace (0-100 scale)
v48, v48_rec	perception of contact with out-groupers at the workplace
	Extent of indirect out-group contact
v43	extent of out-group friends among strong in-group ties

tend to think positively about friends of friends (Wright et al, 1997). To measure indirect out-group contact, respondents were also asked what proportion of their close friends' networks are (not) native Germans, on the same scale used in similar questions: no one, less than half, about half, more than half and all.

B.2) Contact mediators and moderators

Ingroup identification

Social Identity Theory suggests that people strive for a positive sense of social identity. In doing so, they will attempt to distinguish between in-groups and out-groups. This can lead to negative attitudes about out-groups.

In-group identification is measured in the DivCon survey simply by asking the extent to which participants identify with their country (whichever one they had previously indicated as being theirs) on a five-point scale ranging from not at all, coded as one, to very strongly, coded as five. The same question was then asked about how strongly they identified with Europe.

Social identity complexity

Social identity complexity is a concept that refers to the perceived overlap between the groups that an individual is a member of. Three items in the DivCon survey address this notion of social identity complexity. The first of these involves reading out a statement suggesting that the values of the respondent's country are based on their religion. The extent to which respondents agreed with this statement were put on a five-point scale. They were then asked if being a national of that country is the same thing as being a member of their religion, again respondents indicated the extent to which they agreed on a scale of one to five. Finally, participants were asked to estimate the share of Germans that have the same religion as them. Their answer was given as a percentage and it was emphasized that there is no correct answer.

Empathy for foreigners

Empathy and perspective-taking has been shown to decrease both conscious and unconscious stereotyping, as well as increase the overlap between perceptions of the self and of that particular out-group (Galinsky and Moskovitz, 2010).

Empathy is measured in the DivCon survey by reading out four statements and having participants express the extent to which they agree with it on a five-point scale. The four statements include: feeling sympathetic toward foreigners who are discriminated against; caring about the problems of foreigners; being easily able to see things from a foreigner's perspective; and striving to see things from a foreigner's perspective. An empathy scale has been constructed, with a Cronbach's alpha of 0.67, by

combining participants' reactions to these empathy-for-foreigners statements. This scale has been coded to vary from one to represent no empathy to five, which should represent full empathy.

Intergroup anxiety

Anxiety is measured by two items in the DivCon survey. Participants are told to imagine a scenario where they are in a group of people who are of different ethnic backgrounds. They then place the extent to which they think they would feel a) anxious and b) uncomfortable on a five-point scale. The anxiety scale construct is a combination of these items (Cronbach's alpha=0.84). The resulting variable ranges from one, which represents low anxiety, to five, which represents high anxiety.

Ingroup norms

By having friends who themselves have out-group friends, or even just generally have favourable attitudes towards out-groups, we can expect the individual to follow suit if he or she believes that this represents what that group's norms are. In-group norms are measured in the DivCon survey by asking, on a five-point scale, how important the respondent's German friends think it is to be friendly to foreigners.

B.2) Contact mediators and moderators	
	Ingroup identification
v25	identification with the nation
v26	identification with Europe
	Social identity complexity
v28, v28_rec	national values are based on religious values
v29, v29_rec	own nationality means the same as own religion
v30, v30_rec	perceived share of own nationality has own religion
	Empathy for foreigners/perspective taking
v31_1, v31_1_rec	feel sympathy for discriminated foreigners
v31_2	don't care about the problems of foreigners
v31_3, v31_3_rec	can see things from a foreigner's perspective
v31_4, v31_4_rec	strive to also see things from a foreigner's perspective
v31scale	empathy for foreigners scale
	Intergroup anxiety
v4901	feel anxious among out-groupers
v4902	feel uncomfortable among out-groupers
v49scale	intergroup anxiety scale
	Ingroup norms
v73, v73_rec	importance of being friendly to foreigners for German friends

B.3) Personal network

The personal network of the individual is of key importance for the DivCon project. The size, density, proportion of out-group members of participants' personal networks, as well as the relevance of the neighbourhood were addressed. An important distinction when looking at networks is that of strong versus weak ties. Granovetter (1973) describes the strength of a given tie as being characterized by a combination of time spent together, emotional intensity, intimacy and reciprocal services.

Both weak and strong ties that make up an individual's network were acknowledged in the DivCon survey. First, strong ties were explained to the respondent to be people with whom they have frequent contact and discuss personal matters. These strong ties do not live in the respondent's household, but are there when the respondent needs help. Weak ties, by contrast, are described to the respondent as acquaintances with whom they have occasional contact - either in person or over the phone. Weak ties are emphasized not to be close friends that respondents discuss personal matters with.

For strong ties, respondents report the actual number of people who were this kind of friend to them. For weak ties, respondents gave a rough number of how many people in their personal network were this kind of acquaintance. The five possible categories for this question were: up to ten people, 11 to 20, 21 to 40, 41 to 80, and 80 or above. This is coded as one to five, with 80 or more being coded as five.

Having established the number of strong ties, the number of these ties who are (not) native Germans was then indicated. As with other questions in the DivCon survey, non-native Germans are defined as either not being from Germany themselves, or having non-German parents. Once again, respondents were asked to indicate the number of their strong ties that were (not) native German. The equivalent question for weak ties gave five options: no out-group tie, less than half, about half, more than half, and all of them.

The density of the participant's circle of friends has been coded as one for the most disparate category - if the members of the network do not know each other - through to five for the most dense, if everyone in the network does know each other. The focus is then shifted to the non-native German part of the respondent's personal network. On a similar five-point density scale, they are asked if the non-native Germans know the rest of their circle of friends.

Homophily is a central principle in the formation of network ties. We measured the extent of heterophilous ties in the respondents' networks with respect to social class, age, political views and religious beliefs. Five answer options (no tie, less than

half, about half, more than half, and all of them) for each question make these variables comparable to the shares of out-group ties among strong and weak ties which are also measures of homophily.

The characteristics of the out-group members of the respondent's network are then asked about. First, the countries that either the contact, or contact's parents originate from are indicated from a list of possible countries. They were also asked about whether or not some of them had come to Germany as refugees, as ethnic Germans (*Aussiedler*) and if they had been in the country for more than three years. Each of these variables is coded as dummy variables in the data set.

Two questions in the survey touch upon the relevance of the neighbourhood for the out-group ties. First, the proportion of these friends or acquaintances that currently live in the same neighbourhood as the participant is recorded. The possible categories for this question are none, less than half, about half, more than half or all. Using the same coding scheme, they are also asked what proportion of that network they met in the neighbourhood.

Further, respondents were given a list of occasions that they have met their out-group friends or acquaintances. This list includes: work, school or university (one category); an association; another organisation or group; a religious group; while going out; through friends or family; in a neighbourhood they previously lived in; or a different occasion. Each of these categories is represented by a dummy variable.

B.3) Personal network	
	General characteristics of the personal network
v6501	number of strong ties
v6501_rec	number of strong ties in groups
v6502	number of weak ties
v6601, v6601_rec	personal network density
v6602, v6602_rec	density of migrant and non-migrant parts of the network
v7101	heterophilous ties regarding social class
v7102	heterophilous ties regarding age
v7103	heterophilous ties regarding political views
v7104	heterophilous ties regarding religious beliefs
	Immigration-related characteristics of the personal network
v70_XX_rec	166 variables with a specific country of origin
num_v70	personal network: number of countries of origin
v7201	refugees in network
v7202	ethnic German immigrants (Aussiedler) in network

v7203	people who have been living in Germany for a short time in network
	Out-group-related characteristics of the personal network
v6503, v6503_rec	number of out-group strong ties
v6503_rec2	number of out-group strong ties in groups
v6505	share of out-group ties amomg strong ties
v6505_rec	share of out-group ties among strong ties in groups
v6504	share of out-group ties among weak ties
v6504_rec	share of out-group ties among weak ties in groups
v67, v67_rec	share of ties living in the neighbourhood among out-group ties
v68, v68_rec	share of ties met in the neighbourhood among out-group ties
v69_1, v69_1_rec	workplace, school, uni as contact occasion
v69_2, v69_2_rec	association as contact occasion
v69_3, v69_3_rec	another organisation or group as contact occasion
v69_4, v69_4_rec	religious group as contact occasion
v69_5, v69_5_rec	going out as contact occasion
v69_6, v69_6_rec	friends or family members as contact occasion
v69_7, v69_7_rec	a former neighbourhood as contact occasion
v69_8, v69_8_rec	another occasion

C) Outcomes

C.1) Interpersonal trust

As part of interpersonal trust, generalized trust refers to the trust that people have for people in general, regardless of whether or not the individual knows them. Generalized interpersonal trust was measured in this survey by asking participants to indicate the extent to which they agreed with the statement: *Generally speaking, would you say that people can be trusted or that you can't be too careful in dealing with people?* on a five point Likert Scale. This is generally thought to be more effective than dichotomous "agree/disagree" answering options.

In addition to generalized interpersonal trust, measures of interpersonal trust for specific groups were also measured. This was done using the same statement used in the previous question, but for the own group and for four specific groups rather than people in general. These groups are Germans, Western Europeans, *Russlanddeutsche*, and Turks.

C.1) Interpersonal trust	
v5301	trust in people
v5302	trust in Germans
v5303, v5303_rec	trust in own nationals
v5401	trust in Turks living in Germany
v5402	trust in Russlanddeutsche living in Germany
v5403	trust in Western Europeans living in Germany
C.2) Feelings towards specific groups	
v27	feelings towards Germans
v33, v33_rec	feelings towards own nationals
v41_1	feelings towards Turks living in Germany
v41_2	feelings towards Russlanddeutsche living in Germany
v41_3	feelings towards Western Europeans living in Germany

C.2) Feelings towards specific groups

In addition to trust measures, we have also included thermometer feelings measures, whereby respondents are asked to express how warmly they feel about a particular group of people on a scale of 0 to 100. In this survey, respondents were asked about their feelings toward a number of specific groups, again Germans, Western Europeans, *Russlanddeutsche,* and Turks.

C.3) Attitudes to diversity

Three questions in our survey allow us to measure attitudes to diversity. As with the many of the previous questions, interviewers read out statements that touch upon the issue of diversity. The first statement is that it is enriching for a city to have people from different cultures and backgrounds. The second statement touches upon the issue of minority rights, here with reference to the building of mosques in the neighbourhood. Respondents indicated the extent to which they agreed with the statements on a five-point scale. These two items have been combined to construct a diversity beliefs scale, with a Cronbach's alpha of 0.62.

A third question relates to people's attitudes towards the language skills of foreigners living in Germany. Respondents are asked if they feel that these language skills or lack thereof make living together difficult or if they are good enough for general everyday communication.

C.3) Attitudes to diversity	
v50_1, v50_1_rec	diversity is enriching for a city
v50_2, v50_2_rec	right to build mosques, including own neighbourhood
v50scale	diversity beliefs scale
v51	German language skills of foreigners living in Germany
C.4) Attitudes towards foreigners	
v61_1	foreigners threaten the German way of life
v61_2	values of the foreigners are incompatible with the values of Germans
v61_3	foreigners make it more difficult for Germans to find jobs
v61_4	foreigners are a burden on the social welfare system
v61scale	attitudes toward foreigners scale

C.4) Attitudes towards foreigners

Attitudes towards foreigners are often measured in terms of a perceived threat, which is then often separated into two categories: symbolic threat and realistic threat (McLaren, 2003). Symbolic threat focuses more on the perceived threat posed by minority groups on the majority group's culture. Realistic threat refers more to the competition for resources, e.g. jobs, housing or social benefits.

Both symbolic and realistic threats are measured in the DivCon survey, once again with interviewers reading out a number of statements and respondents were asked the extent to which they agreed with it on a five point Likert scale. The statements that correspond to symbolic threat refer to non-Germans threatening the host country's values and general way of life. For realistic threat, the statements include the idea that immigrants take jobs that Germans could be doing and that immigrants are a burden on the welfare state.

These items combine to produce the attitudes towards foreigners scale, with a relatively high Cronbach's alpha of 0.79. This construct varies from zero, for the most negative attitudes to five for most positive.

C.5) Individual and collective efficacy

Both individual and collective efficacy refers "to the capacity for achieving an intended effect" (Sampson et al, 1999: 612-3). These concepts are situational, meaning that an individual or a neighbourhood must be efficacious for a particular task, rather than in a global or general sense.

The DivCon survey includes two questions regarding the impending closure of a hypothetical popular park in the respondent's neighbourhood. This example was chosen as a case for which we assume that a wide range of residents would feel affected and oppose the measure. The respondent is asked: a) what their own action would be in such a situation; and b) the likelihood, on a four-point scale, that the population of the area would protest. For the response of the individual, participants explained what they think they would do, and their answer was noted by the interviewer. We were not interested in the suggested course of action, but in the perceived ability to act. Therefore, answers were later coded into a number of different categories that included: being inactive; suggesting that nothing could be done; taking part in a protest; helping to actually organize a protest.

C.5) Individual and collective efficacy	
v56, v56o, v56_rec	own action against park destruction
v57, v57_rec	residents protest against park destruction
C.6) Political efficacy	
v55_1, v55_1_rec	local politicans represent citizens interests
v55_2	politics is complicated; someone like me doesn't understand
v55_3, v55_3_rec	people like me can influence the local politics

C.6) Political efficacy

The DivCon-survey further includes a number of questions that aim to shed light on the question whether respondents feel politically integrated and represented. Political efficacy is usually differentiated into two dimensions: internal and external. Internal political efficacy refers to the extent to which one thinks that he or she can influence the political process, if they wanted to. External political efficacy refers to beliefs an individual has about the responsiveness of politicians to the concerns of citizens.

Respondents were given three commonly expressed statements that touch upon this notion. A five point Likert Scale was given to report the extent to which people agreed with these statements. The first statement was that politicians represent their interests. Further statements were: that politics is too complicated for people like them to understand; and that people like them can influence the the direction of politics. Between them, these items can measure the extent to which people feel part of political life, feel confident that they can understand politics and can make a difference. Combining these three items only gives us a Cronbach's alpha of 0.33, so these items have not been used to construct a single political efficacy scale.

C.7) Political participation

This variable is captured in the survey with three questions. First, respondents indicated whether or not they voted at the last federal election. Second, non-electoral participation is captured by asking whether or not people have either signed a petition or made a donation for a political issue. Finally, they were asked which party they would vote for if a federal election were held on the following Sunday.

C.7) Political participation	
v58	voted last federal election
v59, v59_rec	support a political issue (petition/donation)
v60, v60s, v60_rec	party vote if federal election next Sunday
C.8) Life satisfaction	
v52, v52_rec	life satisfaction

C.8) Life satisfaction

Finally, life satisfaction is measured on a one to five scale by the conventional question for this variable: "All in all, how satisfied are you with your life?" The possible answers ranged from one for completely satisfied to five for completely unsatisfied.

D) Respondent's background

D.1) Migration background

The DivCon survey includes an extensive section dedicated to the national and migration background of the participant. In the DivCon survey, the respondent's birthplace, parents' birthplaces, citizenship(s), and self-categorized nationality are all recorded.

A common measure of migration background is whether an individual holds citizenship from one or several countries that he or she is not currently residing in. Respondents were asked to indicate all of the countries of which they are a citizen. The number of countries, as an integer, is available in the data set, as well as a dichotomous variable for whether or not they hold citizenship for a country other than Germany. Further to this question is the notion that what is written on your passport differs from the nationality you feel you have. Respondents indicated which country they feel that they belong to, regardless of whether or not they are officially recognized as being a citizen of that country.

To capture immigration, respondents were asked about whether or not they have held German citizenship since birth and what their original citizenship is. They were also asked which country they were born in, and when they moved to Germany, if they were not born there. To capture migration background of a second immigrant generation, the DivCon survey asked in which country the respondent's father and mother was born. If respondents have not held German citizenship from birth, or if they were born abroad, or if one of their parents was not born in Germany, they are coded as having a migration background.

The DivCon survey also asked the respondents whether they regarded themselves as belonging to an ethnic or religious minority. It turned out that this question is not well-understood in the German context.

D.1) Migration background	
v14_XX	46 variables for citizenship
numcitizen	number of citizenships
v14	citizenship(s)
v14_17_rec	non-German citizenship
v15, v15s, v15_rec	nationality (national belonging)
v1401	nationality
v16, v16_rec	German citizenship by birth
v17_XX	48 variables for original citizenship
numorgcitizen	number of original citizenships
v17	original citizenship(s)
v18, v18s, v18_rec	country of birth
v19, v19_rec	living in Germany since the year
residence_g	length of residence in Germany in years
residence_g_grp	length of residence in Germany in 4 groups
v20, v20s, v20_rec	father's country of birth
v20_rec2	non-German-born father
v21, v21s, v21_rec	mother's country of birth
v21_rec2	non-German-born mother
v2101	migration background
v22, v22_rec	reason for migration to Germany
v23, v23o, v23_rec, v23_d	member of an ethnic or religious minority
D.2) Religious background	
v24	religion

D.2) Religious background

For religious background, participants were asked to select from a list of possible religions that was read out to them. The options included Roman Catholic, Protestant, other Christian denomination, Islamic, Jewish, Buddhist, Hindu, other non-Christian, or no religion.

D.3) Socio-demographic status	
v74	year of birth
age	age
age_grp	age in 4 groups
v2, female	gender
v64	household size
v1, hhadult	household size: persons 18+
hhkids	household size: persons under 18
hhkids_d	household with children
v62, v62_rec	partnership
v6301_XX	48 variables for partner's home country
v63	partner's home country
v63_rec	non-German partner

D.3) Socio-demographic status

For socio-demographic status, data regarding the age, gender, number of people living in the household and partner is available. Age is recorded in years and has been recoded into four age groups too. Gender is coded one for female and zero for male. Household size is specified as the number of people usually living in the respondent's household, including people who may be currently absent for reasons such as vacations or being in hospital. Having a long-term partner or not is coded as a dummy variable. The country that the respondent's long-term partner is from is also recorded. A dummy variable for having a non-German partner is available in the data set.

D.4) Socio-economic status

The socio-economic status (SES) variables included in the DivCon survey are education, employment status, occupational status and income. Two questions capture the amount and the type of education that the respondent has completed. The first question asks the highest level of education and then a second question asks about professional qualifications (including university degrees). The data from these two

questions have been combined to construct an education variable that corresponds to the number of years of education completed.

For employment status, various possible categories were included: being in full-time employment, being employed for a few hours a week while also a pensioner or student, or not employed at all. Participants also stated if they were a student, retired, completing military or civilian service, looking after the home or unemployed. They were also asked if they had ever had a full-time or part-time job.

Following up on the type of job that the participant had at the time, or prior to the survey, the various different levels of each kind of job were identified by the participant to allow for the construction of an occupational status variable. This variable ranges from one to five, where five is a higher status, such as senior civil servant or a director of a company.

D.4) Socio-economic status	
	Education
v34, v34s, v34_rec	school education: highest graduation
v35, v35s, v35_rec	highest occupational qualification
edu	education in years
c_edu	education in years (centered around its mean)
	Employment status
v36	employment status
v37	status if not working in a main job
v38	ever held a main job
	Occupational status
v3901	professional group
v3902	type of worker
v3903	type of employee
v3904	type of civil servant
v3905	business owner: number of employees
jobstatus	occupational status
	Income
v75, v75_rec	monthly household income in Euro
v75_rec2	monthly household income in 1,000 Euro
income1	low income until 1,500 Euro
income2	middle income until 3,000 Euro
income3	high income above 3,000 Euro
ln_income	logarithm income

D.5) Extroverted personality

Being extroverted makes people more likely to be able to interact with people, so is usually controlled for when investigating inter-group contact. This index is constructed from three items that capture different aspects of extraversion. Each item involves the interviewer reading out statements such as: liking to have lots of people around; being cheerful and good natured; and enjoying to talk to people. Respondents indicated the extent to which they agreed with these statements on a five point scale. The extrovert personality scale variable has been constructed by combining these three items (Cronbach's alpha: 0.68). This index has been constructed to vary from zero for the least extrovert to five for most extroverted.

D.5) Extroverted personality	
v32_1, v32_1_rec	like having lots of people around me
v32_2, v32_2_rec	a cheerful, good-natured person
v32_3, v32_3_rec	enjoy talking to people
v32scale	extroverted personality scale

E) Interview information

E) Interview information	
	Identifier
intnr	interview id
interv	interviewer id
	Date/time
datum	interview date
tag	interview day
monat	interview month
jahr	interview year
dauer	interview time in minutes
weekday	interview weekday
	Sampling information
herkunft	source of phone number
code_strasse	street id
q649	Kish grid: relative age of respondent
v76	panel consent
	Language
v7601, v7602	interview language

F) Survey design

To avoid biased estimations and statistical results, the rather complex sampling design of cities as primary stage units and areas as secondary stage units should be taken into account. Hence, we created variables that contain information about the survey design (cluster and strata identifiers, population corrections) necessary for STATA's svyset command.

Weights

Three different types of weights are provided by the DivCon dataset. As set out in section 4.5, sampling weights can be used to compensate for differences in the frequency distribution of the sample population, compared to the study population (see Section 2) on key variables to adjust for this difference. This includes sample weights specifically for municipality, sex, age groups, migration background, education and a combined weight variable that takes all of these variables into account. The mzwieght variables act as weights for individuals, whereas the nhdweight weights perform the same function, but for the neighbourhood, to correct context data with regards to age, nationality and sex (separately for each one and a combined nhdweight variable). Finally, cpsweight combines both individual and neighbourhood weights for analyses that deal with multilevel data.

F) Survey design	
	Primary stage units: cities
su1_id	psu cluster identifier: cities (municipality class by diversity by region)
su1_str	psu strata identifier: municipality class by diversity by region (cities)
su1_fpc	psu finite population correction (population size per stratum)
	Secondary stage units: neighbourhoods
su2_id	ssu cluster identifier: neighbourhood (city by diversity by socio-economic background)
su2_str	ssu strata identifier: city by diversity by socio-economic background
su2_fpc	ssu finite population correction (population size per stratum)
	Weights
sampleweight	inverse inclusion probability
mzweight_muc	municipality class MZ2008 weight
mzweight_sex	sex MZ2008 weight
mzweight_nat	nationality MZ2008 weight
mzweight_age	age groups MZ2008 weight

mzweight_mig	migration background MZ2008 weight
mzweight_edu	education MZ2008 weight
mzweight	combined MZ2008 weight (municipality class, sex, nationality, age groups, migration background, education)
nhdweight_sex	sex neighbourhood statistics 2009 weight
nhdweight_nat	nationality neighbourhood statistics 2009 weight
nhdweight_age	age groups neighbourhood statistics 2009 weight
nhdweight	combined neighbourhood statistics 2009 weight (sex, nationality, age groups)
cpsweight	combined poststratification weights (mz, nhd)

G) Contextual data

A number of different contextual data variables have been calculated from the 2009 micro-census and official statistics from the cities included in the survey (see section 6 for more details).

Immigration-related diversity

Although cities collect information on the nationalities of their inhabitants, data protection means that such data are not made available for small areas and every single nationality. We were able to obtain data for four large countries and seven groups of countries: Turkey, Yugoslavia (and its successor states), Italy (incl. enclaves), Poland, other Western Europe, other Eastern Europe, North Africa and Middle East, sub-Saharan Africa, America, Asia (excl. the Middle East), Australia and Oceania. A "missing" category includes the stateless and people of unknown nationality. We derived several variables from this set of data, e.g. number and share of foreigners, share of the three largest foreign nationalities and diversity indices. The visible diversity of the neighbourhood, recorded as low, medium or high, is also available.

Population structure

For each area, we gathered information on population size and density as well as the breakdown into six age groups (0-14, 15-17, 18-24, 25-44, 45-64, and 65 or more years), differentiated by gender (male, female) and nationality (German, non-German). Moreover, categories of age, gender and nationality were combined to form particular sub groups (for example male Germans aged 25-44). However, changes in statutory regulations meant that we were not able to collect context data on changes in population over the past five years.

Socio-economic structure

Available statistics refer to the labour market. We could obtain numbers of unemployed and gainfully employed people.[3] As official figures on unemployment rates are not available for this spatial level, the unemployment ratio was computed by dividing the number of unemployed by all inhabitants between 15 and 64 years. Likewise, we computed the ratio of the gainfully employed at the area level and the city level. Additionally, we calculated a ratio of employees at the city level by dividing the number of employees working in the city by the number of employees living in the city. Unfortunately, it was not possible to collect voter turnout or election data because electoral districts differ from statistical areas.

Urban structure

One important measure of the urban structure is size in terms of spatial extension. Further, we collected numbers of primary and of secondary schools as we assume that schools are contact opportunities in public space. Measures for contact opportunities in public space and residential building structure, both derived from our area explorations, are included.

The variable for contact opportunities in public space ranges from one for few or no contact opportunities to three for multiple opportunities. Area structure value refers to the structure of the buildings in the neighbourhood. Four possible values exist, that vary from dense urban buildings, to areas of detached housing.

G.1) Area context	
code_nb	area identifier
	Immigration-related diversity
foreign_n	share of foreigners in the neighbourhood 2008
foreign2	squared share of foreigners in the neighbourhood 2008
nat_XX_X_09_n	13 variables for number of people with XX nationality (2009 area)
nat_XX_X_09_n_pc	12 variables for share of people with XX nationality (2009 area)
nat_rankXX_09_n	12 variables for number of people from the XX. largest nationality (2009 area)
nat_top3_09_n	number of people from the top 3 nationalities (2009 area)
nat_top3_09_n_pc	share of people from the top 3 nationalities among all foreigners (2009 area)

3 This category includes those included in the obligatory social security schemes.

diversity_f_09_n	diversity index of the 12 nationality groups (2009 area)
diversity_a_09_n	diversity index of the 12 nationality groups + Germans (2009 area)
vdi_n	visible diversity index in the neighbourhood
	Population structure
NAT_GEN_AGE_09_n	63 variables for number of people by nationality by gender by age groups (2009 area)
NAT_GEN_ AGE_09_n_pc	62 variables for shares of people by nationality by gender by age groups (2009 area)
popdensity_09_n	number of people per sq km i.e. population density (2009 area)
	Socio-economic structure
unemploy_n	unemployment rate in the neighbourhood 2008
unemploy_09_n	number of unemployed people (2009 area)
unemploy_09_n_pc	share of unemployed people (2009 area)
sse_lp_09_n	number of gainfully employed, residing in the area (2009 area)
sse_lp_09_n_pc	share of gainfully employed, residing in the area (2009 area)
	Urban structure
prischool_09_n	number of primary schools (2009 area)
secschool_09_n	number of secondary schools (2009 area)
area_n	area in sq km (area)
asv_n	area structure value
pic_n	contact opportunities in public space
G.2) City context	
code_stadt	city identifier
	Immigration-related diversity
nat_XX_X_09_c	13 variables for number of people with XX nationality (2009 city)
nat_XX_X_09_c_pc	12 variables for share of people with XX nationality (2009 city)
nat_rankXX_09_c	12 variables for number of people from the XX. largest nationality (2009 city)
nat_top3_09_c	number of people from the top 3 nationalities (2009 city)
nat_top3_09_c_pc	share of people from the top 3 nationalities among all foreigners (2009 city)
diversity_f_09_c	diversity index of the 12 nationality groups (2009 city)
diversity_a_09_c	diversity index of the 12 nationality groups + Germans (2009 city)

	Population structure
mun_class	municipality class
midtown	mid sized towns (50,000 to 99,999)
bigcity	big cities (100,000 to 499,999)
metcity	metropolitan cities (500,000+)
NAT_GEN_AGE_09_c	63 variables of number for people by nationality by gender by age (2009 city)
NAT_GEN_ AGE_09_c_pc	62 variables for share of people by nationality by gender by age (2009 city)
popdensity_09_c	number of people per sq km i.e. population density (2009 city)
	Socio-economic structure
unemploy_09_c	number of unemployed people (2009 city)
unemploy_09_c_pc	share of unemployed people (2009 city)
sse_lp_09_c	number of gainfully employed, residing in the city (2009 city)
sse_lp_09_c_pc	share of gainfully employed, residing in the city (2009 city)
sse_wp_09_c	number of gainfully employed, working in the city (2009 city)
sse_ratio_09_c	ratio of gainfully employed, working to those residing in the city (2009 city)
	Urban structure
area_c	area in sq km (city)

References

Galinsky, Adam D., and Gordon B. Moskowitz (2000). "Perspective-taking: Decreasing Stereotype Expression, Stereotype Accessibility, and In-group Favoritism". *Journal of Personality and Social Psychology* 78(4): 708-724.

Granovetter, Mark S. (1973). "The Strength of Weak Ties". *American Journal of Sociology* 78(6): 1360-1380.

Laurence, J. (2011). "The Effect of Ethnic Diversity and Community Disadvantage on Social Cohesion: A Multi-Level Analysis of Social Capital and Interethnic Relations in UK Communities". *European Sociological Review* 27(1): 70-89.

McLaren, L. M. (2003). "Anti-Immigrant Prejudice in Europe: Contact, Threat Perception, and Preferences for the Exclusion of Migrants". *Social Forces* 81(3): 909-936.

Sampson, Robert J., and Raudenbush, Stephen W. (1999). "Systematic social observation of public spaces: A new look at disorder in urban neighborhoods". *American Journal of Sociology* 105(3): 603-651.

Statistisches Bundesamt (2009). *Gemeindeverzeichnis*. Gebietsstand 31.12.2008. (Excel sheet).

Winship, C., and L. Radbill (1994). "Sampling weights and regression analysis". *Sociological Methods & Research* 23: 230-257.

Wright, S. C., Aron, A., McLaughlin-Volpe, T., and S. A. Ropp (1997). "The extended contact effect: Knowledge of cross-group friendships and prejudice". *Journal of Personality and Social Pschology* 73: 73-90.

www.bik-gmbh.de/spezialisierung/stichprobensys/cati.html [30/06/2010]

www.randomizer.org/form.htm [30/06/2010]