Ethnic Segregation in Residence, Work, and Free-time
Evidence from Mobile Communication

Ott Toomet[*†‡] Siiri Silm[‡] Erki Saluveer[‡] Tiit Tammaru[‡]
Rein Ahas[‡]

November 13, 2011

Abstract

This paper analyzes ethnic spatial segregation using cellphone data. This allows us to differentiate between place of residence, work, and free-time. We focus on individual pairwise meeting potential (copresence) between ethnic majority and minority groups in a medium-sized bilingual European city (Tallinn, Estonia). We show that segregation in residential and work neighborhoods is rather similar, for both population groups the isolation index ranges between 0.2 and 0.8. However, activities outside of home and work area occur in a virtually non-segregated environment, at least from the spatial point of view. The corresponding isolation index is close to 0.5, the value for that of the random meetings. Our results suggest that physical separation of minorities in segregated neighborhoods may be of less concern than suggested by residential or workplace data only.

1 Introduction

Everyday observations suggest that immigrants are fairly separated from the native population. In particular, political debate is frequently concerned about minorities being concentrated in certain neighborhoods, popularily referred to as "ghettos", where they experience little contact with the majority population. There is a large literature analyzing the potential impact of residential segregation. The main results are ambiguous but a number of studies indicate that physical separation is at least partly causing the inferior labor market and educational outcomes of ethnic and racial minorities (Clark and Drinkwater, 2002; Card and Rothstein, 2007).

However, several authors argue that one cannot easily link residential segregation and integration (Bolt, Özüekren, and Phillips, 2010). As Phillips (2007) stresses, it is hard to "read off" integration from residential segregation. Anecdotal evidence suggest that our most important contacts (besides of the close

[*]Corresponding author. e-mail: otoomet@gmail.com
[2]Department of Economics, Tartu University, Narva 4, Tartu 51009, Estonia
[3]Department of Human Geography, Tartu University

family) are neither neighbors nor colleagues. Similar warnings originate from survey-based and qualitative studies. The linkage between spatial diversity and social network diversity is not trivial (Martinovic, van Tubergen, and Maas, 2009; Peters and de Haan, 2011). Schnell and Yoav (2001) find that dispersion of immigrants out of enclaves will deprive them of their contacts with friends, but fail to create new social ties to the majority population. If true, residential segregation which is the most easily observable form of segregation, might be of a less concern.

Mainly due to data availability, the bulk of the literature focuses on the place of residence and related segregation (see, for instance Cutler, Glaeser, and Vigdor, 1999, 2008). More recently, spreading use of matched employer-employee data has allowed to perform similar analyzes on workplaces (see, for instance Åslund and Skans, 2005; Hellerstein and Neumark, 2008).

Unfortunately, not much is known about segregation in spheres other than residence and work. In particular, very little data is available about social life in free-time. The existing studies, based on friendship ties of school-age children show a substantial social network segregation where members of many minority groups tend to socialize with friends of similar ethnic and racial background (Currarini, Jackson, and Pin, 2009; Martinovic, van Tubergen, and Maas, 2009; Currarini, Jackson, and Pin, 2010).

The current paper complements this literature by analyzing segregation in the area of residence, work, and freetime activities simultaneously, relying on cellphone usage data. We observe the location and time of every call and text message sent in a cellular network in a bilingual European city (Tallinn, Estonia) of about 500 000 inhabitants. Based on the daily activity pattern, we establish the location of home and workplace of cellphone users, employing the anchor point methodology of Ahas, Silm, Järv, Saluveer, and Tiru (2010). In addition, we also observe their preferred language (Estonian or Russian). This allows us to analyze ethnic segregation, based on proximity in space and time, in the area of residence, work, and outside of these two regions. To our knowledge, this is the first paper analyzing these three spheres of activity together. We also control for timing of spatial activities, and in this way we improve over the pure spatial-data based segregation measures. Finally, we add to the rapidly growing literature on analyzing social processes with telecommunication data, using a methodology, comparable to Eagle, Pentland, and Lazer (2009) and Crandall, Backstrom, Cosley, Suri, Huttenlocher, and Kleinberg (2010).

Our analysis indicate that at home and at work, the cellphone users in our data face about similar spatial segregation, with the isolation index ranging from about 0.2 till 0.8. This is true for both language groups. However, outside of home and work regions, the spatial segregation virtually vanishes, with the isolation index being close to that of random meetings. Further analysis indicates that both residential and workplace segregation are positively associated with the freetime segregation, however the effects are small. The results are robust with respect to the choice of spatial units and temporal resolution. Our outcomes suggest that despite of a substantial residential segregation, both minority and majority group members have good chances to meet each other outside their home or work environment.

The paper continues as follows: The next section gives a brief insight to the role of the two dominant ethnic groups in Estonia. Section 3 is devoted to the data description, the following section explains the methodology, including the concept of copresence and homophily, and shows how these measures are computed. Section 5 presents the results with focus on freetime segregation, Section 6 discusses the main findings, and the last section concludes. We leave the robustness analysis and additional results for the Appendix.

2 Background

Before the Second World War, Estonia was ethnically relatively homogenous. By far the largest group were ethnic Estonians (94% of the population of about 1 million). In early years of the war, the country was occupied by Soviet troops and later incorporated into USSR. The years under Stalin's brutal regime thoroughly destroyed the relations with Russians which had been quite friendly up to that time. After the war, as a side effect of industrialization, a steady inflow of mainly Russian-speaking workers from other parts of the Soviet Union moved to the country. This resulted in the population to increase to 1.57 million by 1989, about 40% of which were recent immigrants. A substantial part of the immigrants settled in the capital Tallinn, rendering it to a roughly 50% Russian-speaking city.

The large inflow of Russian-speaking workers combined with the push from Moscow led to increasing importance of the Russian language in the country. Since the 1970s, the country had two *de facto* official languages. The widening use of Russian caused increasing concerns about the future of the Estonian people and the language. One particular outcome of these concerns was an unwillingness to participate in the mainstream Soviet society. Estonians never felt themselves as a part of the Soviet nation and distinguished clearly between their own, "Estonians" and the others, "Russians". In this way these language groupings managed to co-exist in a fairly segregated country.

The country re-gained it's independence on 20 August 1991, during the August Coup in the USSR. The newly elected parliament granted the citizenship only to the citizens of pre-war republic and to their offsprings. Estonian was given the status as the sole official language of the country. These decisions, widely resented by the Russian-speakers, combined with perceived discrimination (Pettai, 2002), historical animosity between the two language groups, segregated school system, and lack of universal bilingualism, have contributed to little inter-ethnic contacts till this day today. The separate worlds are also reflected in media which may present quite different viewpoints depending on the language (Korts and Kõuts, 2002). The relationship between the two main ethnic groups has mostly been "normal" though somewhat tense in periods. Most notably, the tensions exploded to large-scale riots in Tallinn in spring 2007[1].

This historical background renders Tallinn very well suitable for our analysis. First, the population is quite equally divided between Estonian- and Russian speaking groups (54% and 46% respectively, based on the 2000 census). Second, the ethnic composition of the Tallinn neighborhoods is rather diverse (see Figure 1). It is largely an artifact of availability of housing during the peak of immigration in 1970s and 1980s. As there was nothing like a free real estate

[1]The riots were caused by relocation of a Soviet World War II monument, known as the "Bronze Soldier", from central Tallinn to a military cementery. By ethnic Estonians, the monument was considered to glorify the oppressive Soviet rule, while for the Russian-speaking population it was a symbol of victory over nazis in the "Great Patriotic War".

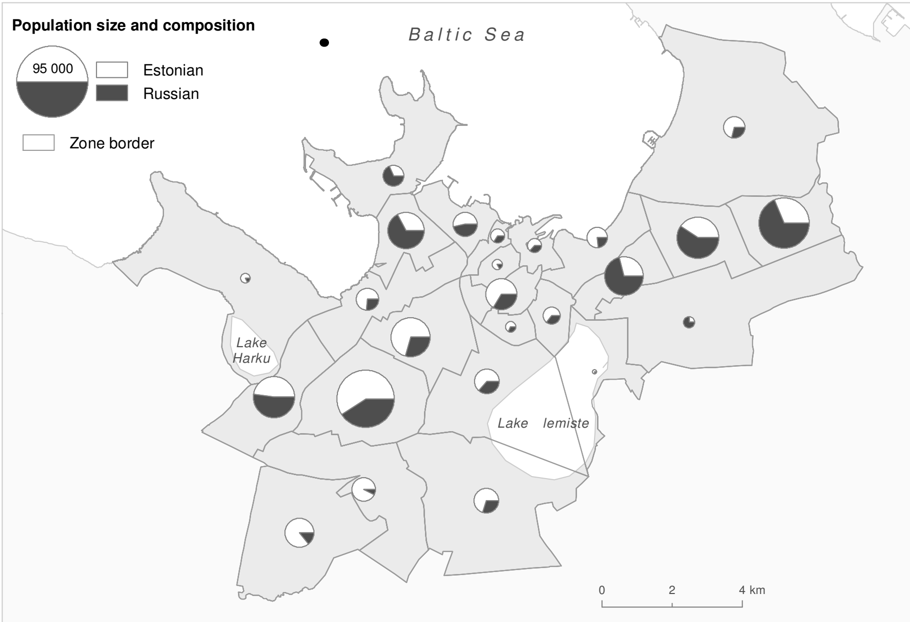


Figure 1: Ethnic composition across Tallinn city tracts

market in Soviet Union, the immigrants had to rely on the official distribution system on flats, mostly in newly built high-rise estates. And third, as bilingualism is not universal, the firms are regularly collecting the information about language preferences on their customers. Our analysis below relies on all these institutional features.

3 Data

3.1 Passive Mobile Positioning Data

This study uses the passive mobile positioning database of the largest mobile service provider in Estonia, EMT. Passive mobile positioning comprises data that are automatically stored in the log files of mobile operators, such as billing information and handover between network cells (Ahas et al, 2008). As the mobile antennas are unevenly distributed, reflecting the population density and transport infrastructure, the accuracy of passive positioning is better in more densely populated areas and around more trafficked roads (Ahas et al, 2008). Based on a 2008 survey, EMT's market share in Tallinn has been estimated at 39%; approximately 96% of the adult population in the country use mobile phones.

The database records the locations of all call activities (outgoing calls and sent text messages), made by mobile phones in the EMT network. The data includes the time of each call (with a 1-second precision) and the mobile antenna through which it was transmitted (cell global identity)[2]. Every network user

[2]"Cell" in case of mobile (cellular) network is the related to one physical transceiver of the

making a call is assigned an identifying tag, making it possible to track the same individuals (or, rather, unique SIM cards).[3] We have no information about the incoming calls, and about the recipients of the outgoing calls.

Besides the calls made in the network, the data also includes some background information. The most crucial of these variables is the *preferred language of the contract holder* which is based on the information, collected by EMT for marketing purposes. Among the users with valid language data, the overwhelming majority uses either Estonian or Russian, and a tiny minority (0.3%) English. We also observe gender and age, associated with most of the phones. Additionally, each phone is associated with place of work and residence. These are work-time and home-time anchor points, based on the timing and location of call activities, and obtained using the methodology by Ahas et al (2010).

The data covers one year time span, from January 1st till December 31st, 2009. We selected individuals for whom we have valid language data, either Estonian or Russian.[4] We also required the individuals to have a valid place of residence in Tallinn, and the corresponding anchor point be located within the same cell for at least 8 months. This resulted in a sub-sample of 32 423 individuals. Among this sub-sample, we randomly selected Estonian- and Russian speakers proportionally to the number of corresponding speakers across the city tracts (see below). In the end, the sample contains 5200 individuals, 2784 (54%) Estonian- and 2416 (46%) Russian-speakers.

We base our main analysis below on 25 *city tracts*, a division of the city based on street network (depicted on Figure 1). Outside of the city, we are limiting our spatial resolution with municipalities in the metropolitan area, and with counties elsewhere. The tracts are rather different both in terms of the population size (spreading between 5 and more than 800 observations) and ethnic composition (percentage of ethnic Russians varies from about 7 to 70%). A more detailed description in given is Table 3 in Appendix A.1.

4 Method

4.1 Copresence

Our analysis is fundamentally about proximity. Specifically, it is based on determining whether two individuals have been together in the same *timeframe*, a certain time interval in the same geographic area. We use the term "timeframe" to stress that the individuals must have been "close" to each other both in space and time.

We define *copresence* for two individuals by counting the number of timeframes where they both have been present. This measure is related to the

network. Usually, this corresponds to a particular geographic area, size of which may vary between a few hundred meters in urban environment to many kilometers in sparsely populated regions.

[3]Obviously, the real identity of individuals and real phone numbers cannot be identified using the tag in our data. The collection, storage, and processing of the data obtained using the passive mobile positioning method complies with all European Union requirements regarding the protection of personal data (EC, 2002), approval was also obtained from the Estonian Data Protection Inspectorate.

[4]The small percentage (0.3%) of English speakers were dropped.

potential number of interpersonal meetings between these persons. This is fundamentally a dyadic measure, i.e. it captures the face-to-face meeting potential among each two individuals in the dataset. Needless to say, we cannot determine whether an actual meeting took place. However, we can still analyze the meeting potential between members on different ethnic groups and how does it depend on individual background characteristics, time, and geographic location. Note that a similar method has earlier been used by Crandall, Backstrom, Cosley, Suri, Huttenlocher, and Kleinberg (2010), using geolocation of photos. Eagle, Pentland, and Lazer (2009) also use a closely related method where they measure copresence by bluetooth technology within about 10-meters distance.

We transform the high-resolution spatial and temporal data to timeframes by aggregating the underlying spatial transceiver information to city tracts, and by rounding time to 3-hour intervals (see Appendix C for different specifications). Denote the set of individuals by $\mathbb{I}$. For each individual $i \in \mathbb{I}$, we measure the time t and location l for her cellphone calls. We denote the timeframe of her call k by c_{ik}. The set of all timeframes where the individual made at least one call is denoted by C_i.

We measure copresence p_{ij} between individuals i and j by counting the number of times these two individuals made calls in the same timeframe. Formally,

$$p_{ij} = \sum_k \mathbb{1}(c_{jk} \in C_i), \tag{1}$$

where $\mathbb{1}(\cdot)$ is the indicator function. Hence, we define dyadic copresence as the number of "common timeframes", places and time intervals where both individuals made at least on call. Note that we do not distinguish between making one or more calls in a given timeframe. This is because we are interested in presence in the given place in given point of time, not in communication activity.

We briefly discuss the most important limitations of our copresence measure. Intuitively, as we only observe location of phone calls, positive copresence is sufficient, but not necessary condition for being in the given place in the given point of time. This may create certain bias for groups with different cellphone usage pattern.

Another point to note is that copresence is based on the presence in a given timeframe. We do not take into account eventual presence in the neighboring places and hence may introduce spurious boundary effects into the data. We choose the current approach because it is much simpler methodologically.

Finally, we do not take into account the duration of stay in timeframes, and the corresponding impact on the meeting potential. For instance, an individual driving around in the city may have copresence with many other individuals despite of only a brief presence in the corresponding regions. This problem can be ameliorated by shortening the temporal dimension of timeframe.

4.2 Homophily

Our primary focus is the meeting potential between two ethnic groups – the chances to meet people with different ethnic background in everyday life. We are accordingly interested in the exposure dimension of segregation (see Massey and Denton, 1988). We select *homophily* as the basis of measuring our meeting intensity.[5] Homophily is essentially a version of isolation index, adapted for single individuals. It measures the percentage of individual's own type of contacts among her complete set of contacts. The analysis below treats the copresence p_{ij} as a measure of contact intensity between individuals i and j. Essentially, we are analyzing the isolation index in copresence.

We define homophily as follows. Denote the observed individual language preferences $\lambda_i \in \{ET, RU\}$, where ET denotes Estonian and RU Russian. We observe two types of ties: between individuals of the same language, and between individuals of different languages. Hence we can write homophily for individual i as:

$$h_i = \frac{s_i}{s_i + d_i}, \tag{2}$$

where s denotes the measure of contacts with the same-language individuals, and d that with the different-language ones. s and d are defined through copresence as

$$s_i = \sum_{\substack{j \in \mathbb{I} \\ j \neq i}} p_{ij} \mathbb{1}(\lambda_j = \lambda_i) \qquad \text{and} \qquad d_i = \sum_{\substack{j \in \mathbb{I} \\ j \neq i}} p_{ij} \mathbb{1}(\lambda_j \neq \lambda_i). \tag{3}$$

Intuitively, homophily equals to the precentage of copresence with one's own language individuals out of one's total copresence. In case of random meetings, expected homophily equals the relative size of one's own group in the population as a whole. Homophily, as a relative measure, is not affected by the daily rhythm of cellphones usage, as long as it is identical for both ethnic groups. Note though that similar homophily figures may mask very different numbers of actual meetings.

The simplest interpretation of homophily assumes that the probability of having a social tie between individuals i and j is proportional to the corresponding pairwise copresence p_{ij} (meeting potential), and this probability is independent of language. This is a brave assumption, but it is qualitatively similar to the implicit assumptions, underlying the interpretation of residential- or workplace segregation measures (such as used by Hellerstein, Neumark, and McInerney, 2007; Hellerstein, McInerney, and Neumark, 2008). The interpretation below still remains valid if this assumption is replaced by a more relaxed one, allowing the likelihood of social ties to differ between same- and different language copresence.

Previous literature on the association between copresence and social ties is rather scarce. Using qualitative methods, Peters and de Haan (2011) find that multi-ethnic contacts in public space do not go beyond superficial interaction, in particular they do not lead to cross-ethnic bonds in private sphere. However, even superficial contacts help to expose and share cultural values and create a more positive view of the others. At the same time, Parreñas (2010) stresses that segregation of Filipina migrant entertainers in Japan partly originates from temporal segregation, the fact that daily schedule for the workers in the nightlife industry is very different from that of the bulk of the population.

Based on electronic data, several studies show that copresence is a strong predictor of underlying social ties (Eagle, Pentland, and Lazer, 2009; Crandall, Backstrom, Cosley, Suri, Huttenlocher, and Kleinberg, 2010). Although we have

[5]Homophily is a commonly used measure in multi-component network analysis (see, for instance, Currarini, Jackson, and Pin, 2009, 2010). See also McPherson, Smith-Lovin, and Cook (2001) for a review.

no information on the actual social ties in our data, we argue that copresence is a more precise measure of interaction potential, compared to only location of residence or location of work based figures. Unfortunately, we have no information in current data to asses the predictive power of our copresence measure.

Note that as we select our sample from inhabitants of Tallinn only, we only measure homophily with respect to the other residents of Tallinn. Copresence with people living elsewhere is not reflected in our data. In particular, this may affect copresence measured outside of the city (although living in Tallinn, individuals may work or spend their freetime outside of the city), as in those regions presumably a lot of time is spent together with people from elsewhere.

4.3 Home, Work, and Freeplace

We analyze segregation in three domains, distinguished by spatial location: in the area of residence (R), at work (W) and in freeplace (F). We call the last domain "freeplace" instead of "freetime" to stress that it is based on location, not time. However, note that both home- and work anchor points are (partly) based on timing of calls.

The place of residence is determined using the anchor point algorithm of Ahas, Silm, Järv, Saluveer, and Tiru (2010), see above. For analyzing segregation in R, we exclude all calls made outside that tract. Hence, only meetings where both partners are in their home region count as R-meetings. In this way, R-homophily is based on the meetings between neighbors and we expect it to reflect closely the population composition in the place of residence. We argue it describes meetings in public space, such as shops, schools and parks nearby, and neighbors visiting each other at home.

Work domain, W, is based on the work-time anchor points. When analyzing this sphere, we exclude all calls outside the work tract, in an analogous way as we do with R. The resulting homophily is related to workplace segregation and describe meetings between colleagues at work, and also encounters with other people, working nearby. Note that for individuals who work in the same region where they live, we are not able to distinguish between these two types of calls and count them as both residence and work-related ones. Hence both of these measures may be somewhat contaminated.

Third, we define freeplace (F) calls as those, conducted outside of R and W tracts. Here we only look at calls, made by both partners outside of their corresponding place of work and residence. We consider this being a proxy for communication during common leisure activities, and other activities not connected to home or workplace, such as shopping or visiting a doctor. Note that F does not include encounters, made in the home or work region of any of the peers.

The location information in our data is based on network cells (antennas or transceivers). For every cell, we can identify the corresponding transmission tower, city tract, and administrative district (see Appendix D for more explanation). Accordingly, we have several options for defining the home and work regions, and the geographic area of timeframes. The main analysis, based on city tracts, focuses on meetings in a rather large area (typical dimension of a city tract is 1km). In case of R and W we count activity in such a large area around the corresponding anchor point as home- or work-related activity, while for F we only look at meetings "sufficiently" far from the place of residence

Table 1: Percentage of copresence across different domains and dyad types.

	1	2	3	4	5
domain	Domain by type			Homophily by domain	
dyad type:	ET-ET	ET-RU	RU-RU	ET	RU
H	0.42	0.50	0.58	0.52	0.59
W	0.32	0.26	0.23	0.61	0.52
F	0.25	0.24	0.19	0.57	0.49
total	1.00	1.00	1.00	–	–

and work. In Appendix C we show that the main results remain robust across different space and time resolutions.

4.4 Data Description: Homophily and Ethnic Composition

This section gives some relevant background information on copresence and homophily in our data.

We start with a rough estimate of importance of different domains. Table 1, columns 1-3, provide a split of different type of copresence across these domains. We can see that the area of residence dominates in terms of where one (potentially) meet the others. This is equally true for intraethnic (both between Estonian speakers and between Russian speakers) and interethnic (ET-RU) meetings. However, the number of meetings in the work area and in the freeplace is not lagging too far behind, both making around 25% of the meetings. We can also see that Estonian-speakers are less home-oriented, compared to Russian-speakers. Hence the table suggests that roughly 50% of all the potential meetings in everyday life occur outside the residential area.[6]

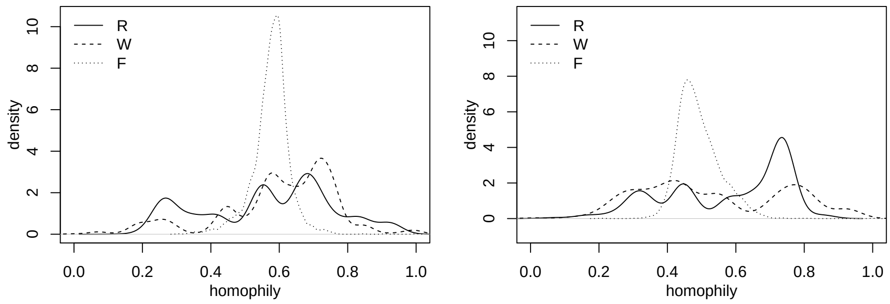
Next, we look at homophily, the percentage of own type of copresence (columns 4 and 5). These figures are roughly between 0.5 and 0.6. The number is slightly higher for Estonian speakers, potentially because they form a somewhat larger group. Estonians appear to be most isolated in work, and least isolated at home, Russians at home and while free, respectively. In general, these average figures are reasonably close to the expected values of 0.54 for Estonian and 0.46 for Russian speakers.

Comparison of R-homophily and corresponding tract population composition shows very close fit (Appendix A.2).

5 Results: F-Homophily

This section focuses on the main results: how closely is homophily in the F-domain related to that in the region of residence and work. We proceed by showing first a series of graphs, and conclude the section with regression analysis.

[6]These figures are sensitive to the definition (size) of the region. See Appendix C.



(a) Estonian (b) Russian

Figure 2: Kernel density estimates of the homophily distribution in R, W and F for Estonian and Russian speaking groups.

5.1 Homophily Distribution in Different Domains

As the first step, we compare the (marginal) distribution functions for homophily in the three domains (Figure 2). In a way the table repeats the columns 4 and 5 in Table 1 but it depicts the distribution functions instead of the mean values. The figure indicates that isolation at place of residence and work are rather similarly distributed, and this is true for both language groups. In both spheres, the isolation index spreads roughly between 0.2 and 0.8, reflecting the different population and workplace composition across the city. The figure also explains why in the table the Estonians' average W-homophily exceeds the average R-homophily while the for the Russian-speakers it is the way around. We can see that a number of Estonian-speakers are living in Russian-dominated areas (with homophily between 0.2 and 0.4) while a significant fraction of Russian-speakers are living in rather Russian tracts (homophily around 0.7). Also the W-homophily of Estonians is more right-skewed than for Russian-speakers.

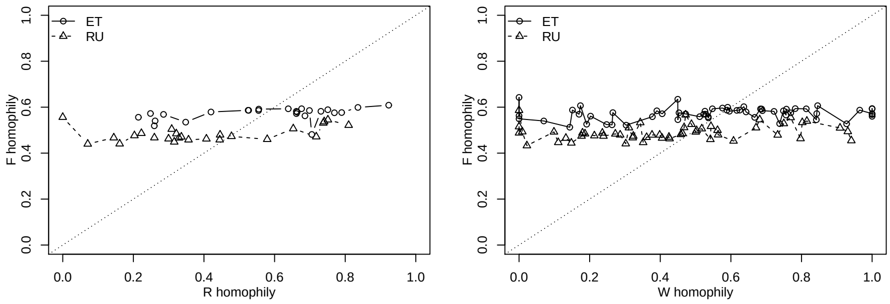
The freeplace homophily, however, is rather different, with virtually all the mass concentrated into a much smaller interval, between about 0.4 and 0.6. The distribution exhibits a prominent single peak for both groups, corresponding to the mean value in Table 1.

This figure clearly indicates that while not at work nor at home, people face rather different and significantly more mixed environment in terms of ethnic composition. Next, we proceed to the 2-dimensional relationship.

5.2 How Closely is F-Homophily Related to the Other Domains?

The previous section presented the marginal homophily distributions for the 3 domains, disregarding their possible interdependence. Here we assesses the interrelationship between freeplace homophily and that in R and W domains. We start the analysis with graphical methods, a statistical approach is given in the section 5.3 below.

We present two figures, depicting the regional average F-homophily as a function of corresponding average R and W-homophily. First, we show the relationship between F and R domains (Figure 3(a)). The average F homophily shows virtually no relationship along the ethnic composition of the place of



(a) Residence

(b) Work

Figure 3: Average homophily in freeplace (vertical axis) by the average homophily in place of residence (left panel) and place of work (right panel). The dotted line depicts the identity relationship.

residence. Outside the home and work environment, the opportunity to meet people of different ethnic background is not closely associated with the ethnic composition of the place of residence. Note that Estonian-speakers show a little larger homophily, as they form a slightly larger group and potentially spend slightly more time in freeplace (or use their mobiles more outside of work and home).

Next, we proceed to the relationship between F and W homophily (Figure 3(b)). As many Tallinn's residents work outside of the city, we have more workplace tracts than residence tracts, and accordingly the workplace homophily values are more outspread. We can clearly see from the figure that W-homophily is not closely associated with F-homophily either.

Finally, we present the geography of freeplace meetings on the map. Most of them occur in the central tracts and downtown-near residential areas, as can be seen from the Figure 4. This is because Tallinn, as a typical European city, is centered around a dense and vibrant downtown, which serves both as the central business district and also the main focal point for cultural activities and entertainment.

5.3 Regression Approach

Here our main objective is to clarify the degree to which segregation in other domains explains the F-segregation. This section complements the graphical analysis above which failed to reveal any strong relationship.

We split the explanatory R and W-homophily into two components: a region specific (macro-level) effect, and an impact of individual deviation from the regional average (micro effect). We specify the macro effects in two ways. First, we control for the average regional homophily $\bar{h}$. This allows us to estimate the association between individual homophily and local macro-level homophily. Second, we introduce region fixed effects instead. Now we cannot identify the macro-effect, but estimated micro effects may be cleaner of macro level mea-

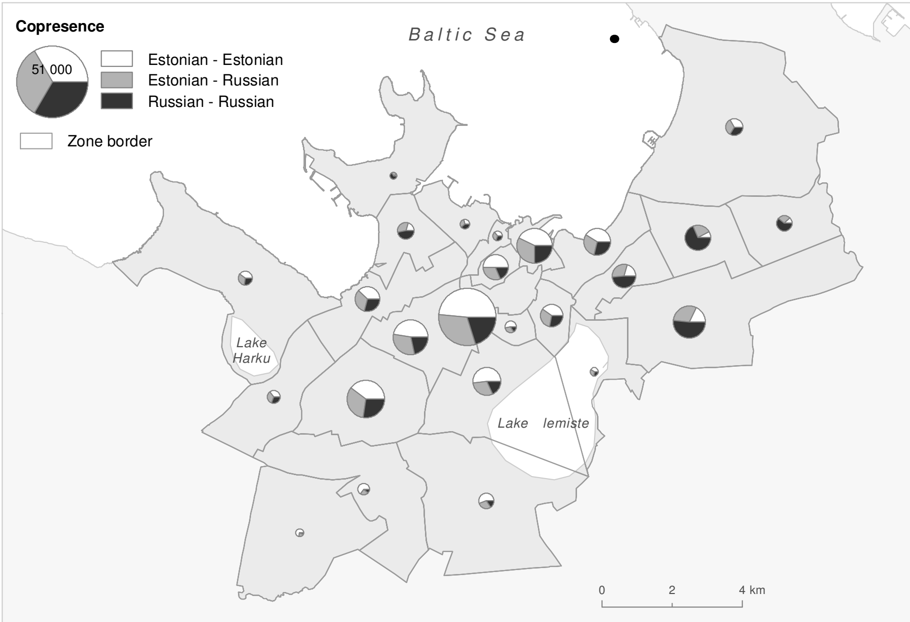


Figure 4: Copresence by location of freetime

surement and specification problems. We have the following two models:

$$h_i^F = \alpha_0 + \tilde{\alpha}_1 \bar{h}^R_{R_i} + \alpha_1 \rho_i + \tilde{\alpha}_2 \bar{h}^W_{W_i} + \alpha_2 \omega_i + \boldsymbol{\beta}' \boldsymbol{X}_i + \epsilon_i \tag{4a}$$

$$h_i^F = \alpha_0 + \tilde{\boldsymbol{\alpha}}_1 \boldsymbol{R}_i + \alpha_1 \rho_i + \tilde{\boldsymbol{\alpha}}_2 \boldsymbol{W}_i + \alpha_2 \omega_i + \boldsymbol{\beta}' \boldsymbol{X}_i + \epsilon_i \tag{4b}$$

Here h_i^F is the F-homophily of individual i; $\bar{h}^R_{R_i}$ and $\bar{h}^W_{W_i}$ are the average R-homophily and W-homophily in the home- and work region of individual i, denoted respectively by R_i and W_i; and ρ_i and ω_i are corresponding individual deviations from the regional average. In fixed-effects version of the model we introduce the residence- and work region fixed effect vectors $\boldsymbol{R}$ and $\boldsymbol{W}$ respectively. $\boldsymbol{X}$ are the individual background variables. We observe three characteristics: age, gender, and call activity which we describe by individual position in distribution quintiles. α and $\boldsymbol{\beta}$ are estimated parameters. We choose to standardize the homophily measures among the explanatory variables ($\bar{h}^R$, $\bar{h}^W$, ρ and ω) to make the results more easily interpretable.

The estimates of both specifications are in Table 2. The table is split into 3 pairs of columns containing three specification for both Estonian and Russian speakers. We can see that h^R, ρ and $\bar{h}^W_W$ are in fact significantly related to F-homophily in all specification, this was not not visible in the graphical analysis above. The estimates of individual deviation in work region, ω, are small and in general not significant. The estimates are rather stable across the specifications and similar for both Estonian and Russian speakers. However, the effects are small. One standard deviation increase in R-homophily is related to an increase in F-homophily by no more than 0.019 units (slightly less for Estonian speakers). None of the estimates for the other explanatory homophily variables exceed these numbers. This leaves only fairly limited room for the other domains to impact homophily in freeplace and explains why we could not

Table 2: Regression estimates of F-homophily

	1		2		3	
	Estonian	Russian	Estonian	Russian	Estonian	Russian
Dependent variable: h^F						
$\bar{h}_R^R$	0.013***	0.018***	0.015***	0.019***		
	0.001	*0.001*	*0.001*	*0.001*		
ρ	0.005***	0.003***	0.005***	0.003**	0.005***	0.004***
	0.001	*0.001*	*0.001*	*0.001*	*0.001*	*0.001*
$\bar{h}_W^W$	0.011***	0.013***	0.011***	0.012***		
	0.001	*0.001*	*0.001*	*0.002*		
ω	0.001	0.001	0.001	0.003*	0.012**	0.004
	0.001	*0.001*	*0.001*	*0.002*	*0.006*	*0.008*
male			-0.014***	0.006**		
			0.002	*0.003*		
age -20			0.003	-0.019*		
			0.008	*0.012*		
age 20-30			0.008***	0.004		
			0.003	*0.004*		
age 55-			-0.004	-0.004		
			0.003	*0.003*		
usage quintile 2			-0.005	0.001		
			0.003	*0.004*		
usage quintile 3			-0.005	0.002		
			0.003	*0.004*		
usage quintile 4			-0.010***	-0.002		
			0.003	*0.004*		
usage quintile 5			-0.015***	0.002		
			0.004	*0.004*		
R^2	0.1733	0.2095	0.2093	0.2265	0.2799	0.3724
# obs	2631	2245	1996	1670	2631	2245
constant	√	√	√	√	√	√
R fix. ef.					√	√
W fix. ef.					√	√

Note: standard errors are clustered across work and home regions.

see these relationships in the figures above. For instance, if we increase the isolation by one standard deviation, both in the residence and work region at the same time, the corresponding F-homophily will increase by about 0.03 (for both Estonian and Russian speakers). Despite of high level of statistical significance, these estimates are unlikely to possess any substantial social meaning. Note though, that we do not know how space-time segregation translates to social segregation, and whether a small change in isolation index can carry important implication in certain cases.

The other explanatory variables reveal little interesting results. Frequent cellphone users and men seem to be slightly less isolated (only Estonians) while Russian-speaking men are a little more isolated. The effects, however, are small.

6 Discussion

The previous analysis gives a clear and unambiguous picture. While the residential and work domains are fairly segregated, the places of freetime activities are not. For various reasons, people of different ethnic origin are living and working in largely separated areas. However, when neither at home nor at work, these individuals have good chances to meet each other, typically in the central districts of the city. This outcome is not entirely surprising and strongly suggests that spatial segregation may be a considerably smaller problem than suggested by residence-only data. But this conclusion may be premature.

First, we do not know if the potential meetings we analyze transform to social contacts of any meaningful value. It is encouraging to find that people who have few chances to meet each other at home or at workplace, can encounter on regular basis in downtown. Unfortunately there exists very little evidence on whether this kind of casual meetings in home, work and leisure domains ever create lasting social ties and how they influence our values. Based on interviews, Peters and de Haan (2011) find that the casual contacts in cityspace around the residential neighborhood remain superficial and seldom exceed simple greetings or brief informative communication. A study of immigrants in Tel-Aviv indicates that exposure to the non-immigrant population at the area of residence does not translate to social ties to the locals (Schnell and Yoav, 2001). These two studies suggest that occasional meetings between strangers, even neighbors, is not sufficient to create more lasting social ties. The weak link between residential and social segregation suggests that these may be leisure time or work-related meetings which carry more weight in shaping individual social networks (see also Boschman, 2011). Unfortunately we are not aware of any related analysis.

Second, nearly perfect integration in freeplace domain in our data is related to the importance of the central business district and other downtown areas. This is presumably because Tallinn is rather compact (a typical size of the city is about 15km) and hence easily accessible. Even more, Tallinn possesses a single, compact and vibrant center which is a natural focal point of various different activities. Our result may not hold for a large, sprawled, or multi-centric city.

7 Conclusions

This paper analyzes ethnic space-time segregation in place of residence, work, and freetime. We use a novel dataset of mobile communication from a bilingual European city (Tallinn, Estonia). The data includes time and location (network cell) of mobile calls and text messages, and preferred language of cellphone owners. By far the most users prefer either Estonian or Russian, as these two ethnic groups form the majority of population in Tallinn.

The analysis is based on copresence, proximity of individuals in space and time. This method improves on the widely used residential segregation indices as it also includes the information on timing. Next, we calculate individual homophily, an isolation index measuring the percentage of own-language copresence in total individual copresence. Based on home and worktime anchor points (Ahas, Silm, Järv, Saluveer, and Tiru, 2010), we find homophily separately for the region (city tract) of residence, work, and the rest of the city (freeplace).

As expected, the homophily in the residence region follows closely the city tract ethnic composition and typically ranges between 0.2 and 0.8. Homophily in the area of work is distributed in a roughly similar way. In contrast, homophily distribution outside of home and work area is considerably more concentrated around the expected value for random meetings (close to 0.5). It is not closely related to the average homophily in the home and work region either. Wherever in the city people live and work, their chances to meet speakers of different language in freeplace remains nearly the same.

These outcomes suggest that residential segregation may be an issue of less concern. Even those who live and work in rather segregated areas tend to spend a substantial amount of time in little segregated cityspace. In Tallinn, this is mainly the downtown area with a large concentration of services and entertainment opportunities. We admit that our encouraging results may only apply to similar compact and monocentric cities.

Our results call for more studies on the association between social ties and geographic proximity. Although we show that the population is reasonably well integrated in the city center, we don't know whether these meetings translate to more or less social ties of any lasting value, compared to encounters close to home or work.

A Data Description

A.1 Number of Observations

Table 3 presents the number of observations and ethnic composition in each tract. It also gives two percentages, one based on the 2001 census data (Pct_{census}) and the other from the current sample. As the data is selected in accordance with the census, both percentages are remarkably similar.

Table 3: Number of observation, and percentage of ethnic Russians by city tracts.

tract	N_{EE}	N_{RU}	pct	Census
Tiskre - Kakumäe - Haabersti	19	4	17.39	16.45
Mõigu	4	1	20.00	11.28
Väike-Õismäe - Astangu	204	223	52.22	52.18
Pelgulinn - Mustjõe	92	33	26.40	26.37
Pelguranna - Sitsi	108	225	67.57	67.49
Nõmme (Laagri - Pääsküla - Kivimäe)	177	30	14.49	14.56
Nõmme (Hiiu - Nõmme)	131	11	7.75	7.58
Mustamäe	488	335	40.70	40.71
Lilleküla	268	110	29.10	29.07
Järve - Tondi - Kitseküla	99	57	36.54	36.42
Nõmme (Männiku - Rahumäe)	107	45	29.61	29.63
Balti jaam	31	16	34.04	34.60
Vanalinn	23	5	17.86	16.45
Kesklinn	161	82	33.74	33.61
Juhkentali	50	28	35.90	35.49
Sadama	30	18	37.50	37.72
Kadriorg	79	24	23.30	23.29
Pirita	85	35	29.17	28.88
Lasnamäe (Mustakivi - Seli)	201	436	68.45	68.48
Lasnamäe tööstus (Ülemiste - Sõjamäe - Väo)	8	24	75.00	75.37
Lasnamäe (Laagna)	176	255	59.16	59.19
Lasnamäe (Sikupilli - Pae)	109	269	71.16	71.07
Kalamaja - Karjamaa	78	68	46.58	46.81
Kopli - Paljassaare	35	75	68.18	68.01
Veerenni	19	9	32.14	32.00
total	2782	2418	46.50	

The resulting gender distribution differs slightly from that in the 2001 census: according to the latter, 45% of residents of Tallinn are male and 55% female. This is true for both Estonian- and Russian speakers. In our sample, the respective figures are 40% and 60% for Estonian-speakers and 47% and 53% for Russian-speakers. The sample also under-represents the youngest (0-19) and oldest (> 60) age groups. This bias is similar across both language groups, and possibly related to different preferences for mobile service providers.

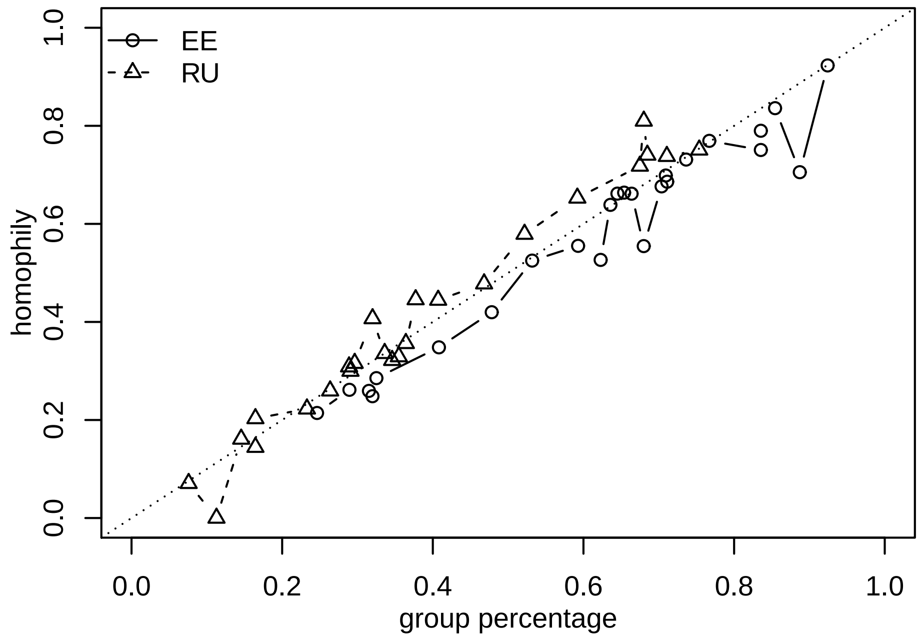


Figure 5: Average homophily in place of residence (vertical axis) by city tract ethnic composition in 2000 census (horizontal axis). timeframe defined as city tract. Dotted line depicts the identity relationship.

A.2 Homophily and Population Composition

Here we split these aggregates by city tracts and analyze the correspondence between R-homophily and the tract ethnic composition. As we select the sample based on the census percentages, we expect the R results to closely reflect the corresponding census figures as the value of homophily in case of random meetings equals to the group percentage. This is indeed the case. Figure 5 clearly shows that these two measures are quite close to each other. It trivially indicates that the average meeting potential in the region of residence is almost perfectly determined by the corresponding ethnic composition. The observed differences from the perfect equality are related to differences in daily schedule, and in some cases to low number of observation. In terms of regression, the regional population composition explains more than 98% of variation in R-homophily.

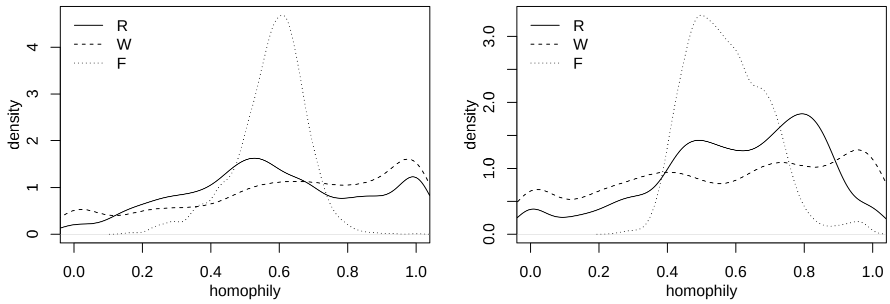
One can easily observe that in most tracts the homophily of Russian-speakers exceeds that of the Estonian-speakers, possibly because ethnic Russians are slightly more prone to stay in the place of residence during the working hours. In that case those who are in their place of residence are rather more together with other Russian-speakers and less with Estonian-speakers.

	Factor loadings	
	Estonian	Russina
Residence	0.611	0.678
Work	0.464	0.453
Freetime	0.567	0.599
% variance	0.303	0.341

Table 4: Factor loadings for the main factor

B Explaining Homophily in R, W and F by a Single Common Factor

In this section we analyze to which degree the homophily in all three dimensions can be explained by a single underlying factor. The results are given in Table 4. The factor loadings are rather similar at about 0.5 for all three spheres. The factor is able to explain about 30% of the total variance. If we split the data between Estonian and Russian speakers, the results for both subgroups are rather similar.



(a) Estonian (b) Russian

Figure 6: Kernel density estimates of the homophily distribution in R, W and F for Estonian and Russian speaking groups.

C Robustness Analysis

In this section we repeat the analysis of Section 5 using different definition of H, W and timeframe. We show that the main results remain robust with respect to the choice of spatial units and temporal length of the timeframe.

C.1 Network Cells

We start with a look at the most fine-grained spatial resolution we have access to, that of the network cells. We also substantially shorten the temporal span of the timeframe, down to one hour. These adjustments radically lower the amount of copresence we observe in our data because the chances to be together in the same network cell in 1 hour time span are much lower than in a city tract during 3 hours.

First, we present the density estimates, analogues to those in Figure 2 (Figure 6). One can easily see that as in the case of larger timeframes, the home and workplace homophily distribution is spread out rather more widely than for freeplace. Note also that because of a much smaller number of observations in each cell compared to the city tracts, all the distributions are spread out substantially more. However, the main conclusion remains the same – in freeplace, people face substantially less segregation than at home or work.

Next, we present the analogues of the figures 3(a) and 3(b). Figure 7 presents the relationship between F and R-homophily, and and 8 that of W-homophily. The main message from both of the figures does not differ from that of the main paper. We see that F-homophily is only weakly increasing in residential and work region segregation. This analysis suggests accordingly that our results are not an artifact of too coarse spatial or temporal resolution.

Finally, we estimate the model of Section 5.3 on the more fine-grained data. Thre results are given in the table 5. For brevity, we only present the estimates of homophily-related coefficients, using similarly standardized coefficients as in the main analysis above. The table indicates that the coefficients are still small, but considerably larger than in case of city tracts. If we were to increase the homophily by one standard error both in the cell of residence and work, the corresponding F-homophily would rise by about 0.06 to 0.07. This is why we are able to see a slight positive slope on the corresponding figures. This is also

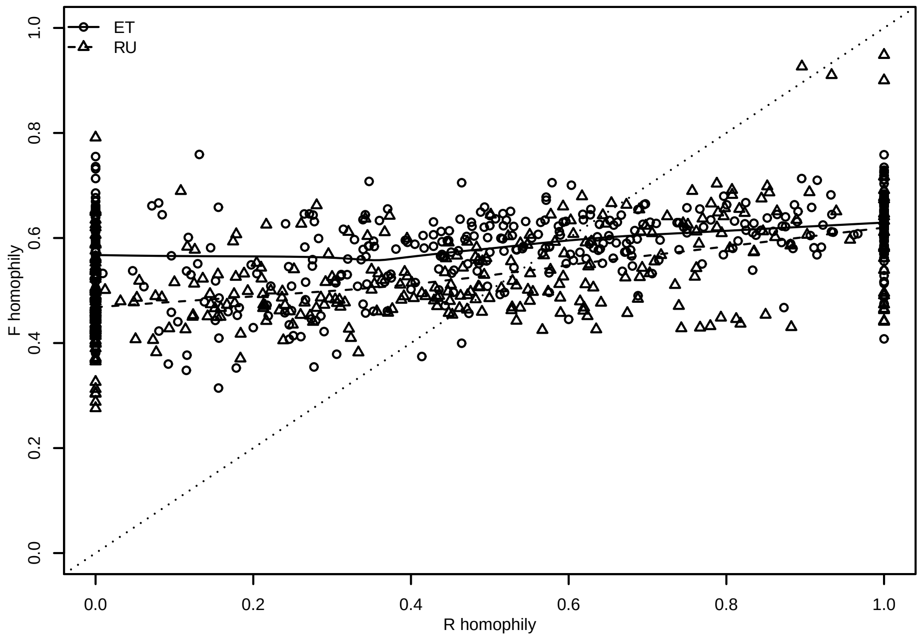


Figure 7: Average homophily in freeplace (vertical axis) by the average homophily in the place of residence (horizontal axis). Solid and dashed lines are the smoothed averages for Estonian and Russian speakers, dotted line depicts the identity relationship.

related to the fact, that the central peak in F-homophily is significantly wider in the current case.

Finally, Table 6 presents a similar factor analysis for cell-based data.

C.2 District-Based Results

In this section we present the results, based on R, W, and timeframe defined as the city administrative district.

First, we illustrate the results graphically (Figure 9). The figure clearly suggest that our homophily measure in R follows closely the corresponding census-based figure (upper left panel). We can also see that work region related homophily is clearly, albeit less closely, related to the residential segregation (upper right). However, the results for freetime (lower panel) indicate virtually no associationship. This suggests that individuals face roughly equal potential of meeting others from different language groups, regardless of the ethnic composition of their area of residence.

Next, we formalize the graphical analysis above by running corresponding OLS regressions for every graph (Table 7). We allow the results to differ by gender and age group. The regression models basically confirm the impression from the figures. We see that homophily in the residence area follows the corresponding census-based measure virtually one-to-one. The corresponding coefficient for the workplace region is about 0.5, meaning that those who live in an area with 10%-points higher percentage of the same group members, tend to

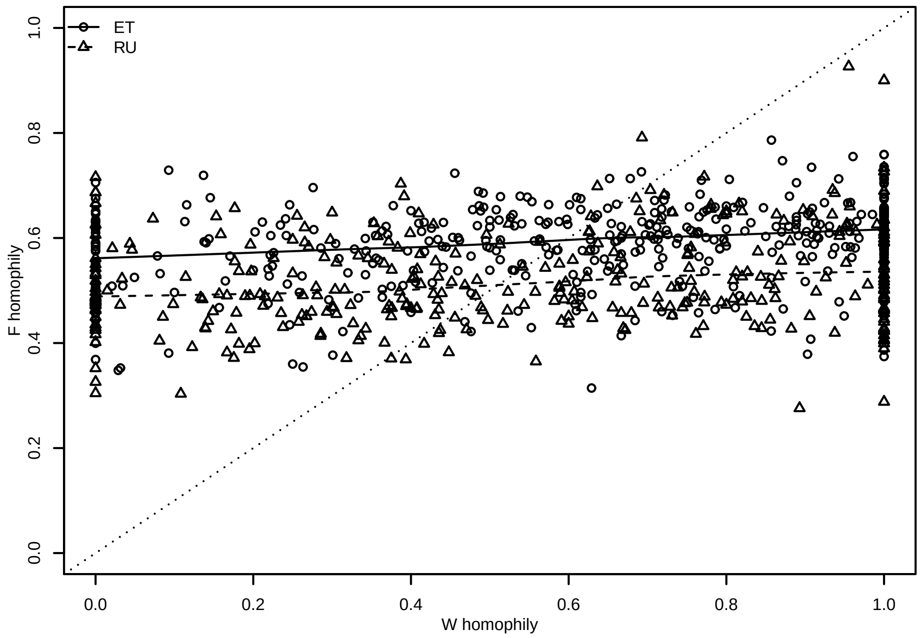


Figure 8: Average homophily in freeplace (vertical axis) by the average homophily in the place of work (horizontal axis). Solid and dashed lines are the smoothed averages for Estonian and Russian speakers, dotted line depicts the identity relationship.

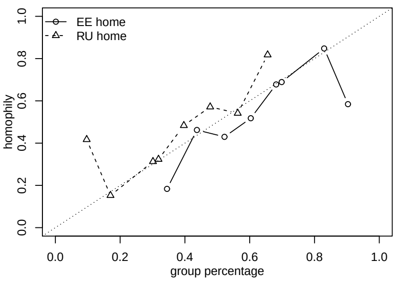
Table 5: Regression estimates of F-homophily

	1		2		3	
	Estonian	Russian	Estonian	Russian	Estonian	Russian
Dependent variable: h^F						
$\bar{h}_R^R$	0.032***	0.038***	0.033***	0.039***		
	0.002	*0.003*	*0.003*	*0.003*		
ρ	0.002	0.001	0.002	0.000	0.003**	0.002
	0.002	*0.002*	*0.002*	*0.002*	*0.002*	*0.002*
$\bar{h}_W^W$	0.030***	0.034***	0.030***	0.028***		
	0.002	*0.003*	*0.003*	*0.003*		
ω	0.024***	0.031***	0.022***	0.024***	0.004	0.000
	0.002	*0.002*	*0.002*	*0.002*	*0.007*	*0.008*
R^2	0.297	0.350	0.323	0.355	0.747	0.795
# obs	2334	2073	1757	1535	2334	2073
constant	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$
indiv charact.			$\surd$	$\surd$		
R fix. ef.					$\surd$	$\surd$
W fix. ef.					$\surd$	$\surd$

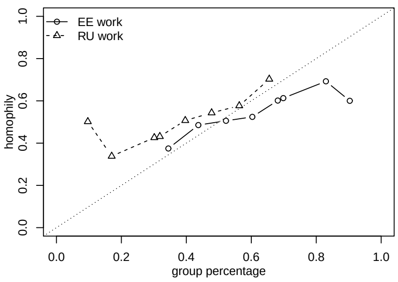
Note: standard errors are clustered across work and home regions.

	Factor loadings	
	Estonian	Russina
Residence	0.459	0.493
Work	0.324	0.384
Freetime	0.997	0.997
% variance	0.437	0.462

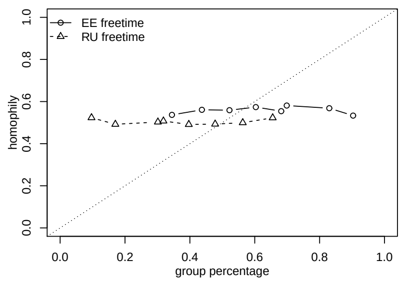
Table 6: Factor loadings for the main factor



(a) Residence



(b) Work



(c) Freetime

Figure 9: Homophily in residence, work, and freetime regions, defined together with timeframe as city districts.

work in a district with 5%-points more members of the same group, in average.

The individual characteristics show little interesting effects. Several demographic groups show somewhat stronger or weaker relationship, but it is hard to draw any conclusion on this pattern.

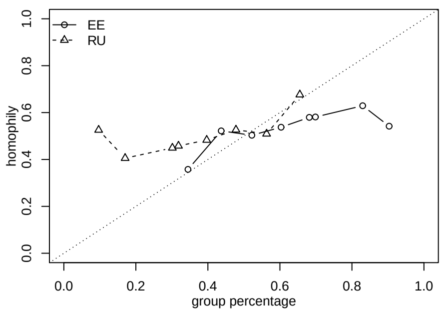
C.3 Transmitter-Based Results

In this section we define R and W based on transmission towers, and timeframe on administrative district. We look at RR, WW and FF homophily (Figure 10).

We see a picture, were homophily depends on the language composition of the residential district substantially more than in case of district-based selection (Figure 9(c)). However, the segregation in the meeting potential outside the home is still substantially smaller than predicted purely based on the residential segregation. This conclusion is confirmed by the corresponding regression results (Table 8). The partial correlation between freetime homophily and the group percentage is 0.336 (for Estonian-speakers) and 0.436 (Russian speakers).

Table 7: OLS estimates of homophily as a function of group percentage (pct). R, W and timeframe based on administrative districts.

Estimate	Residence		Work		Freetime	
	Est	Rus	Est	Rus	Est	Rus
(Intercept)	-0.052***	0.093***	0.280***	0.340***	0.537***	0.496***
	0.007	*0.005*	*0.011*	*0.010*	*0.005*	*0.003*
pct	0.977***	1.035***	0.428***	0.497***	0.026***	0.026***
	0.011	*0.009*	*0.017*	*0.018*	*0.008*	*0.006*
female	0.009	-0.008	0.006	-0.023**	0.002	0.000
	0.008	*0.006*	*0.012*	*0.011*	*0.006*	*0.004*
age -20	-0.016	-0.011	-0.153***	-0.078***	0.020*	-0.026***
	0.014	*0.010*	*0.022*	*0.020*	*0.011*	*0.006*
age 20-30	-0.032***	-0.045***	0.050***	-0.042**	0.020**	-0.002
	0.010	*0.008*	*0.017*	*0.016*	*0.008*	*0.005*
age 55-	-0.032***	-0.052***	-0.203***	-0.106***	0.006	-0.012***
	0.010	*0.007*	*0.015*	*0.014*	*0.008*	*0.004*
pct:female	-0.011	0.014	0.011	0.037*	0.011	0.005
	0.012	*0.011*	*0.019*	*0.021*	*0.009*	*0.007*
pct:age -20	0.024	0.025	0.238***	0.170***	-0.012	0.043***
	0.021	*0.019*	*0.033*	*0.038*	*0.016*	*0.012*
pct:age 20-30	0.065***	0.087***	-0.069***	0.050*	-0.005	0.009
	0.016	*0.015*	*0.026*	*0.030*	*0.013*	*0.009*
pct:age 55-	0.040***	0.087***	0.294***	0.245***	-0.016	0.017**
	0.015	*0.013*	*0.024*	*0.027*	*0.012*	*0.008*
nObs	9000	9000	8836	8930	8953	8949
rSquared	0.7555	0.8221	0.266	0.2669	0.01674	0.01898



(a) Freetime

Figure 10: Homophily in residence, work, and freetime districts

These figures are slightly lower than the corresponding figures for district-based workplace segregation.

Table 8: Homophily as a function of group percentage. Site based results.

Estimate	Freetime	
	Est	Rus
(Intercept)	0.325***	0.357***
	0.016	*0.014*
pct	0.336***	0.436***
	0.024	*0.027*
maleTRUE	-0.021	-0.019
	0.020	*0.018*
age -20	-0.109*	0.067
	0.060	*0.045*
age 20-30	0.004	0.020
	0.025	*0.026*
age 55-	-0.083***	-0.037*
	0.025	*0.022*
pct:maleTRUE	0.019	0.026
	0.030	*0.034*
pct:age -20	0.211**	-0.112
	0.090	*0.085*
pct:age 20-30	0.008	-0.028
	0.040	*0.047*
pct:age 55-	0.134***	0.068*
	0.038	*0.041*
nObs	1000	1000
rSquared	0.4079	0.4321

D Geographic Division of Tallinn

Spatial resolution (regions, places):

- transmitter, cell
- mast, site, tower
- city tract
- district
- municipality
- county

References

Ahas, R., S. Silm, O. Järv, E. Saluveer, and M. Tiru (2010): "Using Mobile Positioning Data to Model Locations Meaningful to Users of Mobile Phones," *Journal of Urban Technology*, 17(1), 3–27.

Bolt, G., A. S. Özüekren, and D. Phillips (2010): “Linking Integration and Residential Segregation,” *Journal of Ethnic and Migration Studies*, 36(2), 169–186.

Boschman, S. (2011): “Residential Segregation and Interethnic Contanct in the Netherlands,” *Urban Studies*, pp. 1–15, forthcoming.

Card, D., and J. Rothstein (2007): “Racial segregation and the black-white test score gap,” *Journal of Public Economics*, 91(11-12), 2158 – 2184.

Clark, K., and S. Drinkwater (2002): “Enclaves, neighbourhood effects and employment outcomes: Ethnic minorities in England and Wales,” *Journal of Population Economics*, 15, 5–29, 10.1007/PL00003839.

Crandall, D. J., L. Backstrom, D. Cosley, S. Suri, D. Huttenlocher, and J. Kleinberg (2010): “Inferring social ties from geographic coincidences,” *Proceedings of the National Academy of Sciences of the US*, 107(52), 22436–22441.

Currarini, S., M. O. Jackson, and P. Pin (2009): “An Economic Model of Friendship: Homophily, Minorities, and Segregation,” *Econometrica*, 77(4), 1003–1045.

——— (2010): “Identifying the roles of race-based choice and chance in high school friendship network formation,” *Proceedings of the National Academy of Sciences of the US*, forthcoming.

Cutler, D. M., E. L. Glaeser, and J. L. Vigdor (1999): “The Rise and Decline of the American Ghetto,” *Journal of Political Economy*, 107(3), 455–506.

——— (2008): “Is the melting pot still hot? Explaining the resurgence of immigrant segregation,” *Review of Economic Studies*, 90(3), 478–497.

Eagle, N., A. S. Pentland, and D. Lazer (2009): “Inferring friendship network structure by using mobile phone data,” *Proceedings of the National Academy of Sciences*, 106(36), 15274–15278.

Hellerstein, J., D. Neumark, and M. McInerney (2007): “Changes in Workplace Segregation in the United States between 1990 and 2000: Evidence from Matched Employer-Employee Data,” in *The Analysis of Firms and Employees: Quantita- tive and Qualitative Approaches*, ed. by S. Bender, J. Lane, K. Shaw, F. Andersson, and T. von Wachter, chap. 5, pp. 163–195. University of Chicago Press.

Hellerstein, J. K., M. McInerney, and D. Neumark (2008): “Measuring the Importance of Labor Market Networks,” Working Paper 14201, National Bureau of Economic Research.

Hellerstein, J. K., and D. Neumark (2008): “Workplace Segregation in the United States: Race, Ethnicity, and Skill,” *Review of Economics and Statistics*, 90(3), 459–477.

Korts, K., and R. Kõuts (2002): “Inter-ethnic communication in Estonian media (in Estonian),” in Valk (2002), pp. 234–245.

Martinovic, B., F. van Tubergen, and I. Maas (2009): “Dynamics of Interethnic Contact: A Panel Study of Immigrants in the Netherlands,” *European Sociological Review*, 25(3), 303–318.

Massey, D. S., and N. A. Denton (1988): “The Dimensions of Residential Segregation,” *Social Forces*, 67(2), pp. 281–315.

McPherson, M., L. Smith-Lovin, and J. M. Cook (2001): “Birds of a Feather: Homophily in Social Networks,” *Annual Review of Socioloy*, 27, 415–444.

Parreñas, R. S. (2010): “Homeward bound. the circular migration of entertainers between Japan and the Philippines,” *Global Networks*, 10(3), 301–323.

Peters, K., and H. de Haan (2011): “Everyday spaces of inter-ethnic interaction: the meaning of urban public spaces in the Netherlands,” *Leisure/Loisir*, 35(2), 169–190.

Pettai, I. (2002): “Mutual tolerance of Estonians and non-Estonians (in Estonian),” in Valk (2002), pp. 213–233.

Phillips, D. (2007): “Ethnic and Racial Segregation: A Critical Perspective,” *Geography Compass*, 1(5), 1138–1159.

Schnell, I., and B. Yoav (2001): “The Sociospatial Isolation of Agents in Everyday Life Spaces as an Aspect of Segregation,” *Annals of Association of American Geographers*, 91(4), 622–636.

Valk, A. (ed.) (2002): *Estonia and Estonians in a comparative perspective (in Estonian).* Tartu University Press, Tiigi 78, Tartu 50410, Estonia.

Åslund, O., and O. N. Skans (2005): “Will I see you at work? Ethnic workplace segregation in Sweden 1985-2002,” Working Paper 2005:24, IFAU, P.O. Box 513, 751 20 Uppsala, Sweden.

References Ahas R, Aasa A, Roose A, Mark Ü , Silm S, 2008, “Evaluating passive mobile positioning data for tourism surveys: an Estonian case study” Tourism Management 29, 469-486.

EC, 2002, “Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications)” Official Journal of the European Communities L201 37-47, 31 July (European Commission, Brussels).