A Theory of Segregation Measurement

Sofía Correa Daniel Hojman*

This version: October 2020

Latest Version Here

Abstract

This paper proposes a theory of segregation measurement based on the intensity and social diversity of pairwise interactions. In our framework societies are described by a space of locations, a space of social groups, and a distribution of agents across locations and groups. Locations can be schools in a district, residences in a city, or platforms such media outlets, where individuals interact. Social groups can defined by race, socioeconomic status, political ideology, or any other social identity. We axiomatize measures that can be expressed as a weighted sum across pairs of an interaction intensity that depends on locations and value of pairwise interactions that depends on social identities. We prove that the index is proportional to the covariance between spatial and social distances, so that high segregation is associated with a high correlation between location and social proximity. We use our framework to study two segregation phenomena. The first one measures socioeconomic segregation in Chilean schools, showing variation across cities in line with residential segregation and across grades in line with differences in elementary and high schools supply. The second one measures ideological segregation in the consumption of media outlets, for different media platforms -newspapers, radio, TV- for 27 European countries, finding systematic differences in segregation across countries and platforms.

*Sofía Correa: Department of Economics, New York University, scorrea@nyu.edu. Daniel Hojman: Department of Economics, University of Chile, dhojman@fen.uchile.cl. We thank Nicolas Suárez for outstanding research assistance. We thank Debraj Ray, Ennio Stacchetti, and seminar participants at he University of Chile, SAET 2017 Conference for the comments and suggestions. Daniel Hojman thanks financial support from the Centre for Social Conflict and Cohesion Studies (Grant No. CONICYT/ FONDAP/15130009)

1 Introduction

Segregation in different domains remains a pervasive social fact in contemporary societies. The lack of socioeconomic and racial diversity of interactions in schools and neighborhoods, and the exposure to like-minded ideological content can hinder a society's ability to embrace the value of diversity. Social inequality is often times both a cause and a consequence of residential, school and cultural segregation. Further, drawing on a long tradition in the social sciences, recent work in sociology has renewed attention on how barriers across social groups and segregation remains a fundamental barrier to equal opportunities. The resurgence of inequalities as focus in local and global politics and the loss of trust in elites and political representatives perceived disconnected from "main street" also points in this direction. Some of the most relevant social and technological recent changes may also contribute to a renewed interest in segregation. The recent mass migration waves across the world are often times associated with new segregated communities in the places of destiny. In a different domain, the radical changes in the media landscape and new forms in which people access to news and information have also be associated with ideological segregation, a potential driver of the recent political polarization. To the extent that different forms of segregation persist or arise, affecting social cohesion and the ability to construct a basic common ground for life in a democratic community, improving our understanding and measurement of segregation remains essential.

This paper provides a general framework to study segregation in different domains. There is a long tradition of theory to study and measure segregation. The starting point of our theory is shared by many conceptualizations of segregation. The basic idea we aim to capture is that segregation is the lack of interactions between individuals belonging to different social groups. Our framework considers a society of individuals that can differ on two dimensions, a social type that defines their social group and a location (or set of locations) they occupy. Depending on the application social types can be race, a measure of socioeconomic status, ethnicity, nationality, religion, ideology or any mix of social characteristics defining the groups we are interested in. An individual's location could be a home address in the case of residential segregation, a school, or the media outlets visited by the individual to acquire information. For example, if the analyst is interested in racial segregation in schools, the social types are races and locations are schools. If our interest is ideological segregation in media consumption, social types are individuals' political ideologies and each individual is associated with a set of locations corresponding to the media she consumes or visits.

The theory presented in this paper has three building blocks. First, we take pairwise interactions as the basic unit of our measure. In practice, the measures we propose aggregate the contribution of each pairwise interaction to overall segregation. Second, the contribution of each interaction depends on two dimensions: the intensity of the interaction and the social characteristics of the pair involved. On the one hand, the more two individuals have access to each other, the more they meet or encounter, the larger the weight of that particular interaction. Intuitively, individuals who occupy the same location interact more than those occupying different locations. On the other hand, the value of an interaction depends on the social types of the pair interacting. In principle, an interaction between two individuals from the same social group contributes more to segregation than one between those belonging to different groups. In sum, both the intensity and the diversity of the interactions matter.

Our axiomatization allows us to obtain a simple formula in terms of these two dimensions characterizing each interaction. Building on axioms that parallel those of Expected Utility theory, Theorem 1 provides a representation of segregation that is precisely the sum over pairwise interactions of a product between a probability that measures the intensity of an interaction and a social value of the interaction that depends on the social types of the pair.

The second general result of the paper shows that the measure can be interpreted in terms of a covariance of distances: a distance in the space of social types —a social distance— and a distance in the space of locations —a spatial or interactions distance. Specifically, we describe a class of problems, called *linear distance-based* problems, which correspond to those in which the intensity and the diversity of the interaction are linear functions of distances in the space of locations and the space of social types, respectively. In the space of social types, it is always possible to define a distance between social groups. For example, if we are only interested in distinguishing whether two individuals belong to same group or not such as it is normally done in the study of racial segregation, two individuals with the same race are assigned distance 0 while two individuals of different races are assigned distance 1. The distance can be richer if the social types space has an order such as unidimensional ideology space or income levels. In the first case, someone who is to the far right of the spectrum is father away from a left-winger than a moderate. Similarly, it is also natural -perhaps more so- to define a distance in the space of locations. Again, if the analyst is simply interested in distinguishing whether two individuals coincide at one location (e.g. student attending the same school), a discrete distance can be used. In other applications, more sophisticated notions of distance could be useful (e.g. time-distance between residences, ideological distance between media outlets, etc.). Theorem 2 shows that if social value of each interaction is inversely proportional to distance in social types and one can express the intensity of interactions as inversely proportional to the distance in locations, then segregation is proportional to the covariance between the social distance -distance across social groups- and the spatial or location distance. Intuitively, segregation is large when two individuals of the same or proximate social groups are spatially proximate to each other, so that most interactions or encounters occur between individuals who share social characteristics.

We distinguish and obtain characterizations for two cases: with and without an individual resource constraint. This resource constraint can be thought as the overall time an individual devotes to interactions. Intuitively, if individuals have a time constraint, the more interactions they have the smaller the time spent on each interaction. This reflects on the interaction intensity measure. It is shown that if individuals have access to a single location (e.g. school, residence) and have resource constraints, minimal segregation configurations are associated with a homogeneous distribution of social groups across locations. This need not be the case in the absence of resource constraints as in this case scale effects may play a role.

The framework is illustrated with two applications. The first one measures socioeconomic segregation in the school system using Chilean microdata, which has administrative information on the socioeconomic status of the parents of each child in fourth and tenth grade. In this case the social type of each student is defined by the education level of the parents and each student's location is simply the school she attends. We compute the segregation for the 22 largest cities of the country. We show that the measures obtained are highly correlated with the most recent measures of socioeconomic residential segregation for these cities, and that the variation across cities seems to be explained by sensible variables such as differences in the structure of local school supply. It also shown that segregation depends on the grade, with more segregation in elementary school than high school.

Our second application measures segregation in media consumption in 27 European countries, using survey data from Eurobarometer. In this application, agents are characterized by an ideology (their social type) and a set of outlets where they get information from (their location on the space). As agents are allowed to get information from many outlets, their location is a vector for each media platform -radio, TV, and newspapers- rather than a single location. Each component of the vector is associated with a particular outlet. As it is natural, in this case individuals can *meet* in more than one location. We find that there is high correlation between segregation levels across media environments, suggesting that there are some fundamental features, probably more related to the idiosincratic political environment, that is explaining segregation. In addition, we find that, for each media platform there is some correlation between segregation and the number of outlets but this is not a general rule.

The paper contributes to the literature on three different margins. First, our framework is relatively general and can be applied broadly. In most previous theories of segregation measurement, both the social types or groups space and the space of locations is given. For example, in a classical paper such as Duncan & Duncan (1955) social types are race and locations residences (see also Massey & Denton (1988)). By considering a more general set up, most known applications can be accommodated by the framework. Second, an important consequence of a more general framework is that is it allows to explore more general propositions. In concrete, by introducing the notion of distances in the interaction space and the social types' space allows to provide a natural interpretation of segregation as a covariance between spatial and social distances of pairs of individuals. Another issue illuminated by our framework is that the segregation order induced by the measure may coincide with widely used measures such as the Duncan or Atkinson measures under some assumptions but not others. Specifically, for these and other measures, minimal segregation is achieved by a configuration in which interactions in each local community reproduces the distribution of social types in the general population. This is true in our framework under some assumptions such as the existence of a resource or budget constraint for each individual, but may not hold if this assumption is relaxed.[1] Finally, the theoretical flexibility in our framework allows to tackle problems in which individuals may encounter at multiple locations such as the consumption of multiple media outlets, as illustrated by our second application. The rest of the paper is organized as follows. Section 1.1 reviews the literature and its relationship with this work. The basic framework and the axiomatization of our measures is presented in Section 2. The general characterization of our measures is summarized by theorems 1 and 2 in Section 3. Section 4.2 presents our two applications, one on socioeconomic segregation in Chilean schools and the other on the ideological segregation of media consumption in European countries. New questions and extensions of the framework are discussed in the conclusion section.

[1] We explore in detail this issue in a companion paper where we show that, in the absence of capacity constraints, minimal segregation could be achieved by configurations in which some locations are associated with a relatively equal distribution of agents while other locations involve segregated individuals, that is, a share of the most numerous groups in locations with a socially homogeneous population. This configuration can sometimes maximize interactions between a social majority and a social minority group.

1.1 Related Literature

This paper aims to contribute to the vast literature on segregation measurement. Most used segregation indexes are measures of *evenness* and *exposure*. The former, focus on the differential distribution of groups across the city, and includes indexes as Dissimilarity, Gini and Atkinson. The Duncan index, which measures occupational segregation in the labor market, is a special case of a Dissimilarity index. Measures of exposure instead, refer to the potential contact between members of different groups, and the most used is the Isolation index.[2] We see our paper lying mainly in the group of exposure measures, as it captures the level of interactions between people with different social characteristics

The main novelty of this paper is the understanding of social spaces and locations as spaces endowed with a distance. Our way of characterizing both the intensity of an interaction and its contribution to segregation in terms of distances gives flexibility to apply this measurement theory to many different frameworks. To the best of our knowledge, this is the first work proposing this characterization. A paper related to ours is Frankel & Volij (2011), who propose two multi-group indexes for the case of school segregation: the Atkinson Index for cases with a fixed number of social groups, and the Mutual Information index for the general case. A novel feature of their indexes is that in general they do not need weighting among different groups according to their size. As theirs, our index does not need weighting, and it goes further than that by allowing flexible distances between social groups.

We contribute with empirical evidence of both economic segregation in schools, and ideological segregation in media consumption. The literature on economic segregation among schools is surprisingly undeveloped, most likely because of the lack of income data at the student level, and because of the focus on segregation by race (see Reardon & Owens (2014) for a detailed review). In a novel study, Owens et al. (2016) study the evolution of income segregation between schools and school districts in the US. To overcome the lack of data, they use the count of enrolled students who are and are not elegible for free lunch as a proxy for income and obtain that segregation has increased since 1990. With the same purpose in mind, we use the educational level of the parents as a proxy for students' income, which allows us to have a finer measure of income levels. We apply our index to obtain economic segregation across schools in Chile, and show that income segregation between schools is closely related with residential segregation.

The literature on segregation in media consumption has increased in the last decade, as both

[2] A complete review about the most common indexes and their properties can be found in James & Taeuber (1985) and Massey & Denton (1988).

scholars and policy-makers become concerned about how the new technologies may facilitate the creation of tailor-made content and increase selective exposure to like-minded views. For instance, Gentzkow & Shapiro (2011) study ideological segregation for media outlets online and offline using a very rich dataset. They compute segregation by using a dissimilarity index. The main difference between our approach and theirs, is that our index considers the fact that agents can interact on different locations at the same time. Moreover, it allows for a more complete ideological scale. Given that nowadays the span of different news outlets ideologies and specific content is bigger, using measures that allow the analyst to use a finer grid of ideological positions makes a difference. In a more recent paper, Kennedy & Prat (2019) use individual-level survey data on news consumption to analyze the potential influence of media outlets on agents voting decisions. Although the focus of their study is on media power, they analyze the link between socioeconomic inequality and information inequality. Although we do not replicate their analysis in this study, it is an example of how our index, through measuring income segregation in an information acquisition environment, can measure the extent to which income and information are related. Although the effects of segregation over political outcomes are out of the scope of this paper, we refer the reader to related works, such as Stroud (2008), DellaVigna & Kaplan (2007), and Campante & Hojman (2013).

2 A Segregation Model based on Pairwise Interactions

2.1 Basic Framework

In our framework agents are characterized by a pair of variables: a social type and the spatial location they occupy. This pair of characteristics are flexible enough to encompass a large number of applications. An agent's social type could be race, income, education, ideology, ethnicity, religion, or any other characteristic (or combination of characteristics) relevant in defining the social groups the analyst is interested in. If we are interested in racial segregation, the social type is race, for example. If we are interested in socioeconomic segregation, the social type could be any measure of socioeconomic status. An agent's location is also quite flexible. In some applications a location is associated with physical space, such as the school attended by a student in the case of school segregation, or a residence address in a city in the case of residential segregation. In other applications, location could be virtual or inmaterial, such as a website or media outlet.

While social types identify social differentiation, locations identify the space in which agents may or may not encounter and interact with each other. The measures of segregation we propose aim to capture the extent to which agents with different social types -i.e., belonging to different social groups- are distantly located in the space of interactions.

The unit on which we measure segregation is called a *community*. This might be, for instance, a city, a set of schools, or a media platform. A community is characterized by:

(i) A set N of agents;

(ii) A *landscape*, Λ, which is the space in which agents interact with each other; and,

(iii) A space of social types, Σ.

We formalize the idea of a landscape in the following definition.

Definition 1 *A landscape* Λ *is a pair* $< L, Q >$, *where*

(i) L *is a set of locations within a community, and*

(ii) $Q = \{Q_l\}_{l \in L}$ *is a collection of location capacities, with* $Q_l \in \mathbb{R}_+$ *possibly unbounded.*

The space of admissibile landscapes is denoted by $\mathcal{L}$.

The allocation of agents in a landscape is described by a space of individual assingments X. Each agent's assingment is a vector $x^i = (x^i_1, ..., x^i_L) \in X$, which describes the locations in the landscape the agent visits. For instance, in the case of schools, the space of individual assignments is described as $X = \{x \in \{0,1\}^L | \sum_{l \in L} x_l = 1\}$. If schools have some capacity constraints, the space of individual assignments would be a constrained space, i.e.

$$X = \left\{ x^i \in \{0,1\}^L | \sum_{l \in L} x_l = 1 \wedge \sum_{i \in N} x^i_l \leq Q_l \right\}. \tag{2.1}$$

We assume that the set of admissible landscapes includes the possibility of complete segregation, i.e., a profile of location assignments such that for any two agents i and j with different social types, if $x^j_l > 0$ then $x^j_l = 0$.

For simplicity we use N to denote both the set and its cardinality. Each agent $i \in N$ is then characterized by a social type $\sigma_i \in \Sigma$, and an assignment into the landscape, $x^i \in X$, where X is the space of possible individual assignments in landscape Λ. Given this, we denote by N_l to the number of agents in location in $l \in L$, and N_σ to the number of agents with social type $\sigma \in \Sigma$.

To fix ideas, consider the case of racial segregation in schools in a city. The set of N agents corresponds to a set of students in the city. The social space is a set of racial groups $\Sigma =$

$\{\text{black, white, asian,...}\}$, and the landscape Λ is composed of a set of schools $L = \{\text{school } 1, \text{school } 2, ...\}$, and a set of capacities for each school $Q = \{Q_1, Q_2, ...\}$, where the latter corresponds to the maximum number of students that each school can admit. Each student's social type is a race $\sigma^i \in \Sigma$. An individual assignment is a vector $x^i = (x_1^i, ..., x_L^i)$ such that $x_l^i = 1$ if student i attends school l, and zero otherwise.

The building block of our segregation measure is the focus on pairwise interactions. In practice, our measures are based on the aggregation of values of each pairwise interaction, weighted by the a measure of intensity of these interactions. Denote by $\Pi(N)$ to the set of possible pairwise interactions between agents in N, with generic element $\pi = (i, j)$.[3] (we omit N when it is clear from the context). Let Δ denote the space of probability distributions over Π. Each pairwise interaction $\pi = (i, j)$ contributes to the level of segregation through two components:

(i) An *intensity*, describing how likely it is that a pair of agents meet in the landscape, represented by the intensity function $\mu : \Pi \rightarrow \Delta$;

(ii) A *social value*, describing the value of an interaction between agents represented by $\rho : \Pi \rightarrow \mathbb{R}$.

Throughout the paper we focus on intensity functions that are location-based. That is, the intensity of an interaction between two agents -how much access they have to each other- depends on their locations. We interpret the intensity function as a descriptive measure of the likelihood that any two people in a community meet or have access to each other. We note however that the relative importance of a particular interaction could be defined by a normative baseline. To illustrate this issue, consider measuring racial segregation in schools. Implicitly, most segregation measure used for this problem assume that any agents in the same school (more generally, two agents in the same location) are accounted equally by the measure. This does not take into account the fact that within a school students may endogenously sort based on homophily, that is, with a tendency to interact more with students of the same race. For example, given two schools A and B one with 20 students, 2 black, 2 asian, 2 latino and 14 white, suppose that in school A students interact randomly disregarding race while in school B students students only mingle with those of the same race. Existing measures would not distinguish two situations. There might be at least two reasons for this. The first one might be a practical limitation of the data: it seems pointless to distinguish between these two schools if the information regarding

[3] More precisely, $\Pi(N) = \{\pi = (i, j) \in N \times N |\, i \neq j\}$.

detailed interactions or social networks within a school is simply not available. The second reason normative: from the perspective of social cohesion and empathy, it may be valuable for students to have access or contact with a diverse population of students.

The social value of interactions is assumed to be group-based, that is we assume that the value of an interaction between two agents depends on their social types, and not on their individual identities. The term "social value" emphasizes the idea that value in a theory of segregation is not associated to a measure of private productivity but instead to the social diversity of interactions. We formalize these ideas in the following definition.

Definition 2 *Let $\mu : \Pi \to \Delta$ be an intensity function and $\rho : \Pi \to \mathbb{R}$ be a social value of interactions function.*

(i) μ is **location-based** *if for some function $m : \Lambda \to \Lambda$, $\mu(i,j) = m(x^i, x^j)$ for all $(i,j) \in \Pi$.*

(ii) ρ is **group-based** *if for some function $r : \Sigma \to \Sigma$, $\rho(i,j) = r(\sigma^i, \sigma^j)$ for all $(i,j) \in \Pi$.*

As seen shortly, it will prove convenient to use an aggregate version of the intensity function, $\tilde{\mu}_\Sigma : \Sigma \times \Sigma \to \Delta$, defined by

$$\tilde{\mu}_\Sigma(s,s') := \sum_{i:\sigma^i = s} \sum_{j:\sigma^j = s'} \mu(i,j). \tag{2.2}$$

This corresponds to the aggregate fraction of interactions between agents in groups s and s'. When is clear from the context, we will omite the subscript Σ to ease notation.

2.2 Distance-based Measures

In this section we show that both social types and interaction locations can be associated with natural notions of distance in each respective space. The motivation for this is both conceptual and computational. As shown in Section 3, our measures conceptualize segregation as a covariance between a social distance and location distance. Intuitively high segregation can be associated with individuals with the same or proximate social types located in the same or proximate locations.

To introduce the notion of distance, consider first the space of social types. In the case of social groups defined by a category such as race or ethnicity, a trivial notion of distance is given by the discrete metric that assigns distance 0 to those in the same group and distance 1 to those in different groups. Indeed, this notion of distance can always be defined and associated to space of social types Σ. There are other case however, in which the structure of Σ is associated with a natural order and distance. For example, if socioeconomic status is defined by a scale such as income or education years, types can be ordered and, moreover, there is a naturally defined metric space. This is also the case for social types defined on an uni-dimensional ideology line.

The case of the space of locations is analogous. Consider, for instance, the case of school segregation by race. It is usually assumed that students only interact with other children in the same school. In that case it is direct to endow the space of interactions with a discrete metric taking value 0 for children in the same school, and 1 for children in different schools. In other cases, such as residential segregation we can think of different notions of distance, such as a discrete distance indicating whether or not individuals live in the same block or geographic unit, geodesic distance or even walking times between two residences.

There are some situations in which we can give more structure to the problem under study. We say that a problem is *distance-based* if it is possible to characterize both the intensity and the social value of an interaction by some notion of distance. More precisely, a problem is distance-based if both the space of individual assignments in the landscape, and the space of social types, are endowed with a metric. Denote these metrics by $d_x : X \times X \to \mathbb{R}$ and $d_\sigma : \Sigma \times \Sigma \to \mathbb{R}$. If this is the case, the intensity and social value functions from Definition 1 can be defined in terms of distances.

Definition 3 formalizes the notion of linear distance-based problems, that will be our relevant framework to obtain Theorems 2 and 3. For simplicity of notation, we define the distance in the landscape by $d_\Lambda(i,j) = d_x(x^i, x^j)$, and the distance in the space of social types by $d_\Sigma(i,j) = d_\sigma(\sigma^i, \sigma^j)$.

Definition 3 *A distance-based problem is a problem in which* (Σ, d_σ) *and* (X, d_x) *are metric spaces. A distance-based problem is linear if the functions* $\mu : \Pi \to \Delta$ *and* $\rho : \Pi \to \mathbb{R}$ *satisfy (i)* $\mu(i,j) = m_0 - m_1 d_\Lambda(i,j)$ *for some* $m_0, m_1 > 0$*, and (ii)* $\rho(i,j) = r_0 - r_1 d_\Sigma(i,j)$*, for some* $r_0, r_1 > 0$.

2.3 Axioms

In this section we state the main axioms. The first one, *Anonimity*, reflects the fact that a segregation order does not depend on agents' identity but on their social characteristics.

Axiom 1 (Anonimity) *Any permutation of agents that preserves the original social groups does not change segregation.*

This axiom allows us to focus on the functions $\tilde{\rho}$ and $\tilde{\mu}$ instead of ρ and μ. The next two axioms refer to how changes in communities, and combinations of them, affect segregation. We combine communities by combining the aggregate intensity functions that characterize them (see equation 2.2). We formalize this idea in the following definition.

Definition 4 *Let C_1, C_2 be two communities, with associated aggregated intensity functions $\tilde{\mu}^1$ and $\tilde{\mu}^2$, respectively. Then, we define the community $C = \alpha C_1 + (1-\alpha)C_2$ as one with aggregate intensity $\tilde{\mu} = \alpha\tilde{\mu}^1 + (1-\alpha)\tilde{\mu}^2$.*

In Section C.1 in the Appendix, we illustrate these operations through an example for the case of socioeconomic segregation in schools.

Axiom 2 (Continuity) *Let C_1, C_2, C_3 be three communities such that $C_1 \succeq_s C_2 \succeq_s C_3$. Then, there exist α such that $\alpha C_1 + (1-\alpha)C_3 \sim_s C_2$.*

Axiom 3 (Independence) *Let C_1, C_2 be two communities such that $C_1 \succeq_s C_2$. Then, for any C_3 and $\alpha \in [0,1]$, $\alpha C_1 + (1-\alpha)C_3 \succeq_s \alpha C_2 + (1-\alpha)C_3$.*

In addition to these technical axioms, we add some axioms that will allow us to define the distance-based notion of segregation. Spatial Proximity refers to the idea that the probability of two agents meeting in the landscape is decreasing in the distance between them.[4] Social Proximity axiom corresponds to the idea that segregation is decreases with diverse interactions. In other words, the more similar two people are, the more their interaction contributes to segregation.

Axiom 4 (Spatial Proximity) *Let $\pi = (i,j), \pi' = (i',j')$ to be two pairs such that $d_\Lambda(i,j) > d_\Lambda(i',j')$. Then, $\mu(i,j) < \mu(i',j')$.*

Axiom 5 (Social Proximity) *Let $\pi = (i,j), \pi' = (i',j')$ to be two pairs such that $d_\Sigma(i,j) > d_\Sigma(i',j')$. Then, $\rho(i,j) < \rho(i',j')$.*

The basic idea of Social Proximity is that diverse interactions contribute to lower segregation. Integrating different groups of people might have effects over economic outcomes through changes in stereotypes, beliefs about others, and changes in social interactions. There is a vast literature on the effects of social and ethnic diversity over several outcomes.[5] For instance, Alesina &

[4] This is the same intuition present in the axiom *School Division Property* in Frankel and Volij (2011): dividing a school is analogous to increasing the distance between agents.

[5] For a detailed literature review on the topic see Alesina & La Ferrara (2005)

La Ferrara (2000) and Easterly & Levine (1997) show a negative relation between social diversity and public good provision and GDP. However, there is also evidence that goes in the opposite direction. Hong & Page (2001) and Hong & Page (2004) develop a theory of problem-solving, in which social diversity is beneficial since it brings with it a different interpretation to a problem, increasing innovation and the probability of solving it. Alesina & La Ferrara (2005) develop a model in which preference diversity may imply public goods under provision, which might be socially costly. Still, there might be benefits in terms of innovation and creativity which, depending on the development of the group under study, might help to overcome the costs. In addition, there might be differences in terms of how the benefits of diversity present over time. Putnam (2007) shows that social diversity costs in terms of trust in the society are observed only in the short term, but in long term, social diversity is welfare improving.

So far, axioms 4 and 5 only ensure that the intensity of an interaction is a decreasing function of the distance in the landscape, and the value of the interaction is a decreasing function of the distance in the types' space. However, nothing can be said in terms of the curvatures of these functions. One could think of many ways in which these functions can be modeled, and although it is crucial for the analysis of interactions, the suitability of each modelling choice to different studies goes beyond the scope of this paper. In the following axioms, we give a structure to this functions that has a nice interpretation, and it is at the same time convenient for applications.

Axiom 6 (Linear Intensity) *The effect of an additive increase in the spatial distance between two agents over their interaction intensity does not depend on their original distance. More precisely, take a pair* (i, j) *with two possible assignments* (x^i, x^j) *and* (x^i_0, x^j_0) *with intensities* $\mu(i, j)$ *and* $\mu_0(i, j)$*, respectively. Then,* $\mu(i, j) - \mu_0(i, j) = K \cdot |d_x(x^i, x^j) - d_x(x^i_0, x^j_0)|$ *for some constant* $K < 0$.

Axiom 7 (Linear Value) *The effect of an additive increase in the social distance between two agents over their value to segregation does not depend on their original distance. More precisely, take a pair* (i, j) *with two possible social types,* (σ^i, σ^j) *and* (σ^i_0, σ^j_0) *with intensities* $\rho(i, j)$ *and* $\rho_0(i, j)$*, respectively. Then,* $\rho(i, j) - \rho_0(i, j) = K \cdot |d_\sigma(\sigma^i, \sigma^) - d_\Lambda(\sigma^i_0, \sigma^j_0)|$ *for some constant* $K < 0$.

3 Representation Results

In this section we state our main results. For the ease of exposition all the proofs are relegated to Appendix A.

3.1 General Representation Result

We first prove that a segregation order can be represented by a function valuing pairwise interactions, which is the baseline for the measures obtained in sections 3.2 and 3.3.

Theorem 1 *The preference order $\succeq_s$ satisfies axioms 1-3 if and only if such preferences are represented by:*

$$S = \sum_{(\sigma,\sigma')\in\Sigma\times\Sigma} \tilde{\mu}_\Sigma(\sigma,\sigma')\rho(\sigma,\sigma'), \tag{3.1}$$

where $\tilde{\mu}_\Sigma(\sigma,\sigma')$ is defined by equation 2.2.

Although this representation might seem general at first, it gives us intuition of how to understand segregation in a society. The basic idea is intuitive: the distribution of agents in the space determines some probability for each type of interactions, so each city is like a lottery of interaction values. To measure segregation we need to measure the value of these *lotteries* for a given segregation order. In the next section, we analyze the case in which interactions can be characterized in terms of distances.

3.2 A Distance-based Representation

We now consider linear distance-based problems, i.e. those problems in which both the intensity and the social value of an interaction are linear functions of distances in the corresponding spaces (see Definition 3). The linearity of the intensity and value functions allows us to prove a very intuitive result: segregation is proportional to a covariance between social and the spatial distances. The idea is that a society is more segregated if similar people are more likely to meet, i.e. similar social types are closer in the space of interactions, or analogously, if different people are unlikely to meet, i.e. different social types are far from each other in the space of interactions. The more these two distances covary, the more segregated the society is. In the extreme, a completely segregated society would be one in which there is a one to one mapping from the space of social types to the space of interactions.

Theorem 2 *A segregation index satisfies Axioms 1-7 if and only if it is proportional to $S = cov(d_\Lambda, d_\Sigma)$.*

Note that axioms 4 to 7 are only consistent with linear distance-based problems. In particular, we can show that a problem is linear distance-based if and only if it satisfies axioms 4-7. An alternative interpretation of Theorem 2 is that segregation is proportional to the coefficient of a regression of the distance in the space of locations, on the distance in the space of social types. Thus, segregation is a measure of the linear association between both.

3.3 Index Normalization

In its more general form, our index is defined by:

$$S = \frac{1}{\Pi} \sum_{(i,j) \in \Pi} \rho(i,j)\mu(i,j) \tag{3.2}$$

In most of the segregation literature, measures are normalized between 0 and 1, where the first number is associated to a configuration that achieves the minimal segregation and the second one to the maximal. For example, for the Duncan index, which is defined for two social types, the minimal segregation is achieved when the population in each location reproduces the global distribution of types in the community. The maximal segregation is achieved for the situation in which all locations are occupied by individuals of the same social group.

Recent work shows that normalizing a segregation index is not trivial as, the implied restrictions, may be associated with a trade-off. In our case, in line with Expected Utility Theory, it is always possible to consider a linear transformation of our measure that respects the segregation ranking and is associated with a normalization:

$$\hat{S} = \frac{S - S^{min}}{S^{max} - S^{min}} = \frac{cov(d_\Lambda, d_\Sigma) - S^{min}}{S^{max} - S^{min}} \tag{3.3}$$

which by construction is 0 for the minimum value of the index, S^{min}, and 1 for its maximum value, S^{max}. However, computing these numbers for a given application is not necessarily immediate. More precisely, the minimal segregation is not necessarily achieved by a uniform distribution of social types across locations as in the Duncan index. As we illustrate below with an example, the configuration that achieves the minimal segregation is not obvious and is sensitive to assumptions that may seem innocuous at first. This issue is treated in detail in a companion paper. Interestingly, the example and characterization of the minimal segregation shows that our measure is associated with a notion of segregation that, depending on these assumptions, may coincide with traditional measures but it may also differ in a meaningful way. We also show that the associate minimization problem is well behaved and can always be solved numerically.

Implicit or explicitly, any segregation measurement takes into account an interaction environment. For residential segregation this environment is a city where residences are located; for racial segregation in schools, it is schools in a school district or city. In the case of ideological segregation in media consumption, the media outlets or platforms available to consumers are locations where agents can ”meet” or encounter. In each of these and other examples we can identify a set of locations, where people can coincide or not. At the same time, in addition to locations, in each application the interaction environment has a configuration that can affect the distribution of agents across locations. For example, in the case of schools, each school has a capacity limiting the number of students that can use that location. In contrast, capacity constraints may not be relevant in the media outlets example, as the number of agents with access to a given outlet is unlimited in most cases. We identify features of the interaction structure that are largely determined by either market forces or public policies.

Other features of the interaction may depend on the nature of the location assignments or demands. For example, in the case of school and residences it seems natural to consider a unit and indivisible usage, so that there is one unit associated to each agent. In the case of media outlets, it is natural to assume that individuals use several locations. Similarly, if we consider time constraints, the strength of interactions can depend on the number of interactions. We consider these issues in the next section.

When considering a normalization of the index, some remarks are in place. First, the set and capacity of locations in a city is not fixed over time. It can be changed by market forces and regulation. In the case of schools, the supply of public and private education and school sizes is typically regulated by local and country-level policies. In the case of residences, the local supply (and density) is affected by local construction regulation and, possibly, by social housing policies. In the case of media supply, capacity constraints may be irrelevant in many cases (e.g. there are no relevant physical constraints that limit the number of visits a website receives in a day). However, the regulation of ownership concentration and policies that aim to foster a pluralistic “marketplace of ideas” can affect the availability of media outlets in the ideological domain.

Second, the landscape affects the space of individual spatial assignments X and places constraints on the spatial assignment profiles at the societal level. For example, in the case of schools, the space of individual assignments can be described as $X = \{x \in \{0,1\}^L | \sum_{l \in L} x_l = 1\}$, so that $X = X(L)$. At the same time, schools' capacity constraints imply that a profile $\mathbf{x}$ $= (x^1, ..., x^N) \in X^N$ satisfies $\sum_{i \in N} x^i_l \leq Q_l$. Hence, changes in the landscape affect globally both individual and social allocation spaces.

Finally, our motivation to define a landscape is driven by practical considerations, both from an analytical and a public policy perspective. Specifically we need to specify an appropriate normalization criterion for our segregation measures. The underlying normalization criteria of segregation measures are often down-played in the literature. Given the above definition, the residential segregation literature, normalization typically considers the landscape as fixed, that is, for a fixed Λ. In practice, this assumes that the built environment is fixed, so that minimal and maximal segregation are found by varying the profile of location assignments of individuals across locations. The underlying thought experiment is that people of different social groups can be relocated in the same city. With a few exceptions this is normally the case in the school segregation literature as well. From a descriptive perspective, it is reasonable to take the landscape as fixed in the short run and from an analytical perspective this is reasonable if the segregation measures provided are invariant to changes in the landscape. However, neither of these assumptions are obvious.

In fact, market supply and public policy that affect the landscape can have a strong impact on segregation. For example, segregation in a city is strongly influenced by new urban developments. In recent work by one the authors of this paper, it is shown that he evolution of the Santiago -one Latin America's most populated cities- is largely determined by housing policies that resulted in the allocation of poor families in new peripheral neighborhoods of the city -large government-funded projects, new secluded high-income developments in the borders of the city, and densification of central neighborhoods. On the other hand, changes in media technologies and markets have drastically affected the supply of content in the media market, affecting ideological segregation and polarization (see Campante & Hojman (2013) and Levy & Razin (2019)). Landscape can also be affected by regulations and government initiatives. In education, publicly-funded school supply and capacities. In an urban setting, construction regulations, location of social housing, are important examples. In media markets, the regulation of media concentration ownership, entry and also policies to foster balance in the ideologies of media supply provide additional illustrations.

At the same time, from a normative stance, it is far from obvious whether segregation measures should be invariant to changes the landscape. In a companion paper we show for the case of school segregation that the nature of assignments that minimize segregation can vary substantially across different landscapes. Perhaps more importantly. If changes in the landscape have important effects on segregation or the minimum and maximal segregation, why should segregation normalization take them as fixed? To illustrate this issue consider two societies A and B with the same population and population distribution of social groups. Suppose that in A there are 1000 small schools and in B a 10 large schools, why should segregation normalization take the number of schools as fixed if this is a variable that can be changed over time? Similarly, if A and B are two countries one with 1000 news websites and the other with 10, why should we normalize segregation taking the media landscape as given?

In principle, our framework does not take a stance on whether the landscape should vary in the calculation of maximizing segregation. In fact, in the short run the landscape is not flexible. We may consider two variants of the optimization problem leading to the maximal and minimum segregation levels. The first one takes Λ as given and considers the location assignment profile as the optimization variable:

$$\min_{\mathbf{x}\in\chi(\mathbf{\Lambda})} S(\mathbf{x},\sigma;\Lambda). \tag{3.4}$$

The solution of this problem is some profile $\mathbf{x}^*(\mathbf{\Lambda})$. The value of the problem is $\Phi(\Lambda) = S(\mathbf{x}^*(\mathbf{\Lambda},\sigma;\Lambda)$ (analogous for maximization). If minimal and maximal segregation allow to vary the landscape, we can consider a second stage optimization across admissible landscapes,

$$\min_{\Lambda\in\mathcal{L}} \Phi(\Lambda), \tag{3.5}$$

yielding a solution Λ^*. In the Appendix, we describe the normalization optimization problem for the application to ideological segregation in media outlets, which is solved using numerical methods.

In the following section we specialize the measures to the case of unit location assignments and individual capacity constraints. In this case, large class of problems, we provide an explicit normalization and formula.

3.3.1 Individual Capacity Constraints

So far, we have considered the case in which agents do not have capacity constraints in their interactions with others. For example, given two schools A and B, one with 20 students and the other with 100, and assuming that agents interact only with agents in their school, the intensity of interaction functions give equal weight to any interaction in school A and school B. This could be a reasonable assumption for situations in which what matters is whether two students share the same space but not the amount time spend with each other. If empathy were mostly associated with sharing a common location with people of a different race —access to a different social group— but not the time spent with people of different groups, this makes sense. This is also a sensible assumption if most relevant activities in school involve the whole population. In many cases however, it makes sense to assume that each agent has a capacity constraint, a time o resource budget that she allocates to each interaction (see, for example, Echenique & Fryer Jr (2007)).

In this section, we extend the framework to consider agents having a time budget or capacity constraint. In the school case, a capacity constraint implies that the probability of meeting other students decreases with the size of the school, or, in other words, the time spent with each classmate decreases as the number of classmates increases. Thus, the weight of a particular interaction in school A, the smaller school, larger than the weight of an interaction in school B, the larger school.

We introduce a new axiom that formalizes the notion of resource constraints. The intuition is that, as agents have a limited time to interact with each other, to increase the interaction with one agent they have to reallocate their resources across the space, decreasing the intensity of interaction others.

Axiom 8 (Resource Constraint) *Agents have limited resources to interact with each other, which is represented by the following resource constraint:*

$$\sum_j \mu(i,j) = T \qquad \forall i \tag{3.6}$$

We restrict the analysis to a framework in which agents can only be assigned to a unique location. This imposes a restriction over the space of individual assignments X, for any given landscape Λ. The following definition formalizes this notion.

Definition 5 *An individual assignment is a Unit Location Assignment (ULA), if (i) $x_i \in \{0,1\}^L$, and (ii) $x_i^T e_L = 1$.*

Let $\Pi_l = \{(i,j) | x_i = x_j = e_l\}$ to be the set of possible pairwise interactions within location $l \in L$, and denote by $\bar{d}^l_\Sigma = \frac{1}{\Pi_l} \sum_{(i,j)\in\Pi_l} d_\Sigma(i,j)$ to the average social distance between pairs of agents within location l.

Proposition 1 *Consider an assignment satisfying (ULA), and suppose axioms 1-8 hold. Then, the normalized segregation index is given by*

$$S = 1 - \frac{1}{\bar{d}_\Sigma} \sum_{l=1}^{L} w_l \bar{d}^l_\Sigma, \tag{3.7}$$

where $w_l = \frac{N_l}{N}$ is the share of the population in location l.

The proof can be found in the Appendix. Observe that the index can be expressed as

$$S = \sum_{l=1}^{L} w_l \left(1 - \frac{\bar{d}^\Sigma_l}{\bar{d}^\Sigma}\right), \tag{3.8}$$

that is, a weighted sum across locations of an expression that compares the local social distance with the global social distance, with weights equal to the share of the population in each location. This is parallel to the dissimilarity index, which can also be expressed as a weighed sum across locations of a quantity that compares the local share of a minority relative to the aggregate share of the minority.

4 Applications

4.1 A Economic Segregation in Schools

This application illustrates the use of distance-based segregation measures for the case of schools. Given the poor availability of income data at the student level, the empirical work in economic segregation among schools is not very developed. On an attempt to fill this gap, in this paper we use administrative individual-level data of Chilean students provided by the Ministry of Education to explore socioeconomic segregation across cities. For each student in 4th grade and 10th grade in 2014, the dataset identifies the school they attend in addition sociodemographic and family background variables.[6] Although we do not have direct information about students' family income, we use the parents' educational level —average years of education of mother and father— as measure of each student's socioeconomic status. Segregation measures are computed for all regional capitals, including the three Chilean metropolitan areas, Santiago, Valparaíso and Concepción.

We denote each school by $l \in \{1, ..., L\}$, and l_i corresponds to the school attended by student i. We use N to denote the total number of students in a district, N_l to the number of students at school l, and $N_{g,l}$ to the number of students in quintile g attending school l. Also, denote by Π and Π_l the number of pairs in the population, and in school l, respectively.

In this context, the landscape has the form of a partition, which is constitent with unit location assignments (see Definition 5). As students only interact with other students attending their same schools, and assuming that interaction is uniform, the corresponding metric in the space of interactions is a discrete metric.

We divide the income distribution into income quintiles, and associate to each student the

[6] In Chile, students of these grades in all schools —with a few exceptions— are required to take a standardized test in math and language, the SIMCE. In addition to the individual scores, the Ministry uses a complementary questionnaire to gather sociodemographic and family background information. The dataset covers roughly 95 percent of all Chilean schools, excluding new and special education schools.

average parents' educational level of the corresponding quintile. Thus, $d^{\Sigma}(i,j) = |y_i - y_j|$, where y_i is average years of education of the respective quintile. Then, from proposition 1 the normalized index can be computed using the following equation:

$$S = 1 - \frac{1}{\overline{d}_{\Sigma}} \sum_{l \in L} \frac{N_l}{N} \overline{d}^{l}_{\Sigma} \tag{4.1}$$

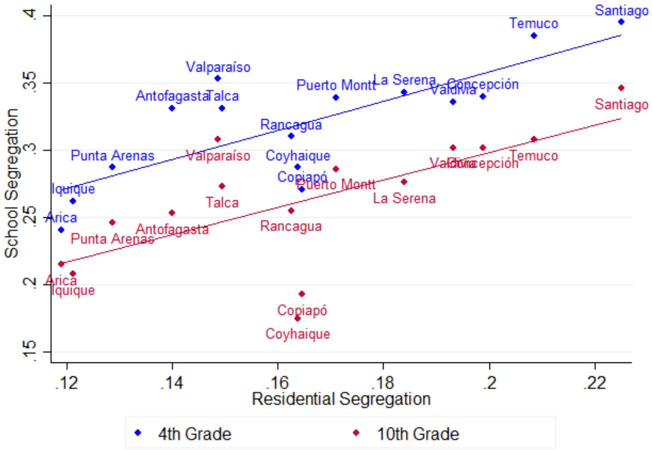
where $\overline{d}_{\Sigma}$ is the average social distance in the population, and $\overline{d}^{l}_{\Sigma}$ is the average social distance at school l.

The results are shown in Table ??. A few remarks are in place. First, for both grades segregation is highly correlated with the size of the city (in terms of number of students and schools). In particular, Santiago is the most segregated city in both grades, followed by Temuco and Valparaiso. Secondly, in each city we have that segregation is higher in 4th grade than in 10th grade. These facts, are consistent with residential socioeconomic segregation patterns in Chilean metropolitan areas and the role of distance in school choice. Indeed, we compare our school segregation index with the residential segregation measures for Chilean cities obtained by Agostini et al. (2016), who use household income as an SES measure and Census data. We obtain a correlation coefficient between school and residential segregation of 0.79 for 4th grade students, vs a 0.67 for 10th grade. Even when both are high, the correlation for 4th grade is slightly higher, probably reflecting the fact that younger children have more mobiliy constraints than students in 10th grade. Both comparisons are plotted in Figure 1.

Table 1: School socioeconomic segregation by city

	4th Grade			10th Grade		
Region	S	Schools	Students	S	Schools	Students
Santiago	0.395	1,244	50,535	0.346	753	41,020
Temuco	0.385	109	3,031	0.308	189	7,320
Valparaiso	0.353	340	9,079	0.308	108	6,145
La Serena	0.343	87	2,773	0.276	31	1,626
Concepcion	0.340	279	9,351	0.302	39	1,522
Puerto Montt	0.339	100	2,886	0.286	44	2,434
Valdivia	0.336	60	1,552	0.302	18	1,119
Talca	0.331	67	2,658	0.273	55	2,146
Antofagasta	0.331	74	4,214	0.253	44	3,107
Rancagua	0.310	76	2,930	0.255	42	2,612
Coyhaique	0.287	25	725	0.175	45	2,906
Punta Arenas	0.287	37	1,409	0.246	35	1,355
Copiapo	0.271	35	1,965	0.193	30	1,553
Iquique	0.262	55	2,176	0.208	11	470
Arica	0.241	65	2,422	0.215	22	946
Mean	0.321	177	6514	0.263	98	5085
Max	0.395	1244	50535	0.346	753	41020
Min	0.241	25	725	0.175	11	470

Figure 1: School vs Residential Segregation.



Source: Residential segregation from Agostini et al. (2016).

4.2 Segregation in Media Consumption

In many situations, agents can interact with people in multiple locations. This is particularly frequent in virtual environments, like the case of media consumption: agents can obtain information from different outlets, and on each of them they *meet* different people. This meeting can take the form of an actual interaction, like in an online news forum, or simply obtaining the same information, like in a newspaper. In this section, we approach this problem by analyzing the problem of ideological segregation in media consumption in a framework in which each location corresponds to a specific media outlet and individuals may consume more than one outlet.

Usual metrics of segregation that are equivalent to using a discrete metric, like the Isolation index (see Gentzkow & Shapiro (2011)), fail to consider two features of the media consumption problem: (i) agents are characterized by a rich set of social characteristics, like ideologies, and this is not captured by a discrete metric over the social space; (ii) agents can obtain information from (and henceforth, interact through) more than one media source. Our framework aims to tackle these issues.

We consider a set up in which agents have a budget of time that can allocate across news sources. The landscape is composed of a set of media outlets, and the *locations* correspond to the time they spend on each of them. We analyze three markets separately, i.e. segregation in TV consumption, newspaper consumption and radio stations. The social characteristic we are interested in is political ideology, and then our segregation index measures the the extent to which individuals with different ideologies are sharing media consumption.

Let $\Lambda = 1, ..., L$ to be the set of media outlets on each market, and $x_l^i \in \{0, 1\}$ an indicator variable taking the value 1 if i consumes outlet $l \in L$ and zero otherwise. As defined in Section 2.1, the vector $x^i = (x_1^i, ..., x_L^i)$ summarizes individual i's consumption bundle, and we focus on the normalized version $\hat{x}_i$ with generic element $\hat{x}_l^i = \frac{x_l^i}{\sum_{l=1}^L x_l^i}$. Individuals interact to the extent the share media consumption, that is, if they coincide in their locations. The effective number of location coincidences between two individuals i and j is then $\sum_{l=1}^L \min\{\hat{x}_l^i, \hat{x}_l^j\}$. This defines our distance in the landscape:

$$d_\Lambda(i, j) = 1 - \sum_{l=1}^{L} \min\{\hat{x}_l^i, \hat{x}_l^j\}. \tag{4.2}$$

To see the intuition of this distance, consider the following example. There are two outlets, $\Lambda = \{1, 2\}$, and two agents i and j. Agent i gets information from both outlets, spending half of the time on each, but agent j only acquires information from outlet 1, spending all the time there. Formally, $\hat{x}^i = (0.5, 0.5)$ and $\hat{x}^j = (1, 0)$. Then, they both share only half of the total time together: they *coincide* $\min\{0.5, 1\} = 0.5$ on outlet 1, and $\min\{0.5, 0\} = 0$ on outlet 2. Their distance is $d_\Lambda(i, j) = 1 - 0.5 = 0.5$. Of course this is an approximation of the actual time they spend on the same outlet, as in our data we only have information about which outlets the agent visits, but not how much time she spends there.

We use $y_i \in \{1, 2..., 10\}$ for the answer of individual i in the ideological self-identification question. We define the social distance between i and j as $d_\Sigma(i, j) = |y_i - y_j|$. The average social distance of a uniform matching pairing is $\bar{d}_\Sigma$, which is computed as in the previous application.

Our index is defined as:

$$S \quad = \sum_{(i,j)\in\Pi} \mu(i, j)\rho(i, j) \tag{4.3}$$

for some linear functions μ and ρ consistent with Definition 2 (see Section 2.2). From theorem 2, we can directly computed as being proportional to the covariance between social and spatial distances, i.e.

$$S = \frac{1}{\Pi} \sum_{(i,j)\in\Pi} d_\Sigma(i, j) d_\Lambda(i, j) - \bar{d}_\Sigma \bar{d}_\Lambda \tag{4.4}$$

We compute a normalized version of the index. The maximal segregation corresponds to a media environment completely segregated, which is one in which all agents sharing the same ideology visit the same outlet. To compute this, we construct such environment (redefining agents consumption in a one-to-one relation with their ideology), and computing equation 4.4. As we argue in Section 3.3, we compute the minimal segregation numerically. In most countries, this corresponds to having some groups fully segregated, and others distributed uniformly across media outlets. A detail of the optimization problem solved can be found in Appendix B.

We use survey data from Eurobarometer 82.4 (2014) that covers 28 European countries. The survey has a series of questions that ask individuals about the TV stations, radio stations, newspapers, and websites they use.[7] We decided to leave websites out of the analysis, as the consumption of websites could be including consumption of radio and newspapers online, and this could generate a biased measure.

The survey also asks individuals to self-identify in an ideology uni-dimensional ten point scale, where 1 is the extreme left and 10 is the extreme right.[8] In Table ?? we show some descriptive statistics of our dataset.

Table 3 shows the index values and number of outlets for each country and media environment. The country with the highest segregation is Malta, which has the maximum value for the index in radio, TV and newspaper market. Cyprus has the lowest index value for both TV and Newspapers, while the Netherlands have the lowest segregation in radio. In half the countries the market of newspapers is the most segregated, while in other 8 countries the TV market is the most segregated.

In general the correlation between segregation and the number of outlets is low but positive, ranging from 0.04 in the radio market, and 0.28 in TV. This is in line with theories predicting that increasing competition in the market for news, understood as increasing the number of competitors, could exacerbate segregation as allow consumers to self-select more effectively into news outlets with like-minded opinions (see Mullainathan & Shleifer (2005)).

What is more interesting is the high correlation observed across media markets. The correlation between segregation in radio and newspapers is of 0.68, while for TV and newspapers is 0.80. This might suggest that there are some structural political conditions that generate this segregation levels. To explore this idea, we compare segregation indices for each environment with an index of polarization obtained from the Varieties of Democracy (V-Dem) Project (see Coppedge (2020)) As we show in Figure 2, segregation and polarization are positively correlated in all media outlets.

[7] Questions QP17a,b,c y d.

[8] Media consumption is captured by questions QP17a,b,c and d, and ideology by question D1.

Figure 2: Polarization and Segregation

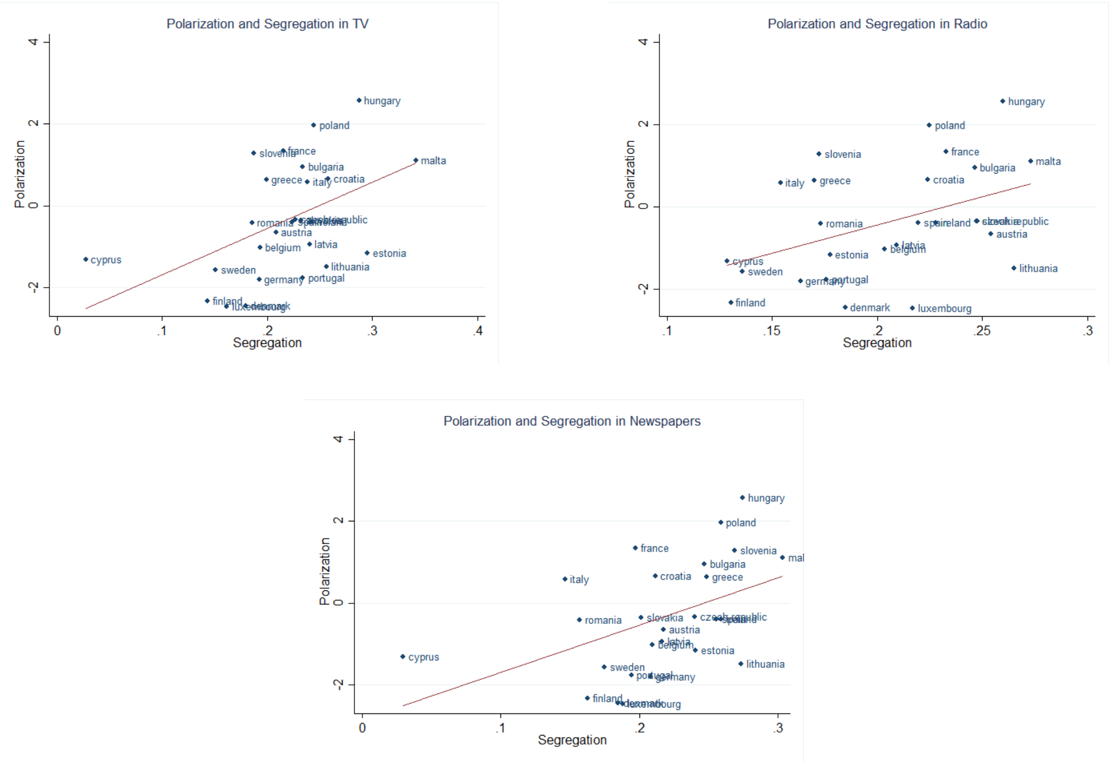


Table 2: Ideological self-identification in european countries, Eurobarometer 2014

Country	**Mean**	**Median**	**Std. Deviation**	**Sample size**
Austria	4.83	5	1.91	929
Belgium	5.08	5	2.00	924
Bulgaria	5.48	5	2.63	784
Cyprus	5.44	5	3.00	288
Czech Republic	5.39	5	2.36	949
Germany	4.89	5	1.75	1381
Denmark	5.47	5	2.35	955
Estonia	5.71	5	2.21	679
Spain	4.34	5	1.91	838
Finland	5.52	5	1.93	893
France	5.01	5	2.16	829
Great Britain	5.06	5	1.84	1092
Greece	5.06	5	2.08	729
Croatia	5.38	5	2.47	786
Hungary	5.84	5	2.32	842
Ireland	5.28	5	1.84	838
Italy	5.05	5	2.18	686
Lithuania	5.17	5	2.63	708
Luxembourg	5.25	5	1.94	411
Latvia	5.88	5	2.15	781
Malta	5.10	5	2.34	303
The Netherlands	5.08	5	1.85	969
Poland	5.94	5	2.42	799
Portugal	4.68	5	1.90	668
Romania	6.01	5	2.82	691
Sweden	5.22	5	2.35	1008
Slovenia	4.60	5	2.61	684
Slovakia	5.12	5	2.37	869

Table 3: Ideological Segregation in european countries, Eurobarometer 2014

	TV		Radio		Newspaper	
Country	$\hat{S}$	L	$\hat{S}$	L	$\hat{S}$	L
Austria	0.208	26	0.254	24	0.218	29
Belgium	0.193	26	0.204	29	0.210	27
Bulgaria	0.234	19	0.246	11	0.247	28
Cyprus	0.028	18	0.129	22	0.030	19
Czech Republic	0.227	23	0.247	18	0.240	29
Germany	0.192	28	0.163	27	0.208	28
Denmark	0.179	13	0.185	9	0.185	29
Estonia	0.295	22	0.177	25	0.241	29
Spain	0.224	23	0.219	22	0.256	28
Finland	0.143	15	0.130	22	0.163	28
France	0.215	27	0.233	22	0.197	27
Great Britain	0.247	26	0.235	24	0.293	28
Greece	0.199	25	0.170	26	0.249	25
Croatia	0.258	24	0.224	26	0.211	26
Hungary	0.287	19	0.260	21	0.275	27
Ireland	0.242	21	0.228	26	0.259	26
Italy	0.238	19	0.154	21	0.146	27
Lithuania	0.256	23	0.265	26	0.274	17
Luxembourg	0.162	5	0.217	12	0.188	25
Latvia	0.240	16	0.209	17	0.217	26
Malta	0.341	18	0.273	25	0.303	21
Netherlands	0.130	19	0.103	28	0.160	29
Poland	0.244	28	0.225	30	0.259	29
Portugal	0.234	24	0.176	23	0.194	29
Romania	0.185	29	0.173	27	0.157	19
Sweden	0.151	9	0.136	7	0.175	25
Slovenia	0.188	18	0.172	22	0.269	19
Slovakia	0.232	22	0.247	24	0.202	26
Min	0.028	5	0.103	7	0.030	17
Max	0.341	29	0.273	30	0.303	29
Average	0.213	21	0.202	22	0.215	26

5 Conclusion

This paper proposes a theory of segregation measurement based on the intensity and social diversity of pairwise interactions. In our framework societies are described by a space of locations and social groups, and a distribution of agents across locations and groups. Both the space of location and the space of social groups are flexible enough to include many different segregation problems. Locations can be schools in a district, residences in a city or platforms such media outlets, where individuals interact. Social groups can defined by race, socioeconomic status, political ideology, or any other social identity.

We axiomatize measures that can be expressed as a weighted sum across pairs of an interaction intensity that depends on locations and value of pairwise interactions that depends on social identities. We prove that the index is proportional to a covariance between spatial and social distances.

We then use our index to study two segregation problems. First we measure socioeconomic segregation in Chilean schools using Chilean microdata, which includes information on socioeconomic status of the parents of each students. There is variation across cities and grades, and school segregation is highly correlated with residential segregation.

As our index allows for multiple simultaneous interactions, in a second application we use it to measure ideological segregation in the consumption of media outlets, for different media platforms -newspapers, radio, TV- for 27 European countries. There are systematic differences in segregation across countries and platforms, suggesting that there are some fundamental features, probably related to the political environment, that explains these segregation levels.

A Appendix: Main Proofs

Proof of Theorem 1 From von Neumann Morgenstern utility representation theorem, a complete and transitive preference relation satisfies continuity and independence if and only if admits a expected utility representation:

$$S(\mu) = \sum_{(i,j)\in\Pi} \mu(i,j)\rho(i,j).$$

Using anonimity axiom we get the result. □

Proof of Theorem 2 First, let's prove that axioms 1-7 imply $S = cov(d_\Lambda, d_\Sigma)$.

It is direct to see that axioms $1-3$ imply: $S = \sum\limits_{(i,j)\in\Pi} \mu(i,j)\rho(i,j)$. By axioms 4 and 5,

$$S = \sum_{(i,j)\in\Pi} f(d_\Lambda(\lambda_i,\lambda_j))g(d_\Sigma(\sigma_i,\sigma_j)) \tag{A.1}$$

for some decreasing functions f, g. Moreover, by axioms 6 and 7, the functions $f(\cdot)$ and $g(\cdot)$ are linear. Then, there exit constants m_0, m_1 and r_0, r_1 such that:

$$S = \sum_{(i,j)\in\Pi} (m_0 - m_1 d_\Lambda(\lambda_i,\lambda_j))(r_0 - r_1 d_\Sigma(\sigma_i,\sigma_j)) \tag{A.2}$$

Note that $\sum_{(i,j)\in\Pi} \mu(i,j) = 1$ imposes a constraint over parameters m_0, m_1, such that:

$$m_0 = \frac{1}{\Pi} + m_1 \overline{d}_\Lambda \tag{A.3}$$

Plugging this in equation A.2 and with a little algebra we obtain the result:

$$\begin{aligned} S &= m_0 r_0 - r_0 m_1 \overline{d}_\Lambda(i,j) - r_1 m_0 \overline{d}_\Sigma + r_1 m_1 \frac{\sum_{(i,j)} d_\Lambda(\lambda_i,\lambda_j) d_\Sigma(\sigma_i,\sigma_j)}{\Pi} && \text{(A.4)} \\ &= \frac{r_0 - r_1 \overline{d}_\Sigma}{\Pi} + r_1 m_1 \left[\frac{\sum_{(i,j)} d_\Lambda(\lambda_i,\lambda_j) d_\Sigma(\sigma_i,\sigma_j)}{\Pi} - \overline{d}_\Lambda \overline{d}_\Sigma \right] && \text{(A.5)} \\ &= \overline{\rho} + r_1 m_1 cov(d_\Lambda, d_\Sigma) && \text{(A.6)} \end{aligned}$$

which completes the proof in this direction.

Now suppose a segregation measure proportional to $S = cov(d_\Lambda, d_\Sigma)$. Note that the contribution of each interaction takes the form: $(\overline{d}_\Lambda - d_\Lambda(\lambda_i,\lambda_j))(\overline{d}_\Sigma - d_\Sigma(\sigma_i,\sigma_j))$, which is a decreasing function of the spatial and social distances. Moreover, any additive variation in $d_\Lambda(\lambda_j,\lambda_j)$ only changes the first component in an additive way, and same for additive variations in $d_\Sigma(\sigma_i,\sigma_j)$. Thus, axioms 4-7 hold. The proof of axioms 1-3 is analogous to the previous theorem. This completes the proof. □

Proof of Proposition 1 In order to prove the proposition, we first prove the following auxiliary lemma.

Lemma 1 *Consider an assignment satisfying (ULA), and suppose axioms 1-8 hold. Then,* $\mu(i,j) = \frac{2}{N(N_l-1)}$ *for any* $(i,j) \in \Pi_l$.

Proof of Lemma 1 We already know that $\mu(i,j) = 0$ if $x_i \neq x_j$. Since $\mu(i,j) = \mu(d_\Lambda(i,j))$, we have that $\mu(i,j) = \mu_l$ for each $(i,j) \in \Pi_l$. With and individual resource constraint, the latter implies that $\mu_l(N_l - 1) = T$, from which $\mu_l = \frac{T}{N_l-1}$. Now, since μ is a probability distribution,

$$\sum_{i,j} \mu(i,j) = \sum_{l=1}^{L} \Pi_l \mu_l,$$

where, with some abuse of notation, $\Pi_l = N_l(N_l-1)/2$ is the number of pairs in l. Using the previous expression for μ_l, we must have that $T = N/2$, which yields that $\mu(i,j) = \frac{2}{N(N_l-1)}$ for any $(i,j) \in \Pi_l$. ■

From theorem 2 and lemma 1, the expression for S becomes

$$S = \frac{2}{N} \sum_{l=1}^{L} (\frac{1}{N_l - 1}) \sum_{(i,j)\in\Pi_l} \rho(i,j) = \sum_{l=1}^{L} w_l \bar{\rho}_l,$$

where $w_l = \frac{N_l}{N}$ is the share of the population in location l, $\bar{\rho}_l = \frac{1}{\Pi_l} \sum_{(i,j)\in\Pi_l} \rho(i,j)$, and $\Pi_l = N_l(N_l-1)$.

Now, with (DB), if $\rho(i,j) = r_0 - r_1 d^\Sigma(i,j)$, with $r_1 > 0$, then $\bar{\rho}_l = r_0 - r_1 \overline{d_l^\Sigma}$. Combining this with nn, we have

$$S = r_0 - r_1 \sum_{l=1}^{L} w_l \overline{d_l^\Sigma}.$$

Note that $S \leq r_0$, which is achieved with equality if and only if $\overline{d_l^\Sigma} = 0$ for all l. This is indeed the case if all agents in each location have the same social type. A sufficient condition to achieve this maximal segregation is thus that $L = G$ is an admissible landscape. Normalizing maximal segregation to 1, it follows that $r_0 = 1$.

On the other hand, minimal segregation is obtained by maximizing the expression $\sum_{l=1}^{L} w_l \overline{d_l^\Sigma}$, as a function of the number of agents of each social type in each location. It can be shown that this es achieved by the population distributed uniformly across locations or by having all agents in the same location (which is equivalent to setting $L = 1$). In this case , for all l, $\overline{d_l^\Sigma} = \overline{d^\Sigma}$. Minimal segregation is thus $S^{min} = r_0 - r_1 \overline{d^\Sigma}$ and normalizing minimal segregation to 0, yields $r_1 = 1/\overline{d^\Sigma}$. □

B Appendix: Minimal Segregation in Media Consumption

Let G be the number of social types. For any two social types $\sigma, \sigma' \in \Sigma$, define $A_{\sigma,\sigma'} = d_\Sigma(\sigma, \sigma') - \overline{d}_\Sigma$.

Suppose assumption ULA holds. LetN_σ be the number of agents with social type σ, and $N_{l\sigma}$ the number of agents type σ in location l. We define $x_l = (x_{1,l}, ..., x_{G,l})' \in \mathbb{N}^G$ as a vector such that $x_{l,\sigma} = N_{l,\sigma}$. In this context, the minimization problem reduces to:

$$\begin{aligned} \min_{x_l \in \mathbb{N}^G} \qquad & \sum_{\sigma \in \Sigma} \sum_{\sigma' \in \Sigma} A_{\sigma,\sigma'} \cdot \sum_{l=1}^{L} N_{l,\sigma} N_{l,\sigma'} \\ \text{s.t.} \qquad & \sum_{l=1}^{L} N_\sigma \end{aligned}$$

This is the general problem to be solved. Note that the problem corresponds to minimizing a quadratic function over a simplex.

C Appendix: Examples and Additional results

C.1 Community Operations: An Illustrative Example

When comparing segregation levels across communities, we keep fixed (N, Σ, ρ), and study how changes in the distribution of agents over the space of interactions affect segregation.

Consider the study of school segregation by income. Students are characterized by their family income, which can be *low*, *middle* or *high*. We denote this social space by $\Sigma = \{l, m, h\}$. The proportion of each social group in the population are given by $\left(\frac{N_l}{N}, \frac{N_m}{N}, \frac{N_h}{N}\right) = (0.25, 0.5, 0.25)$. The space of locations Λ is composed by eight schools, all of which have the same capacity.

Each agent i has associated a social group $\sigma^i \in \Sigma$, and a an assignment $x^i \in X$. We assume that students only interact with other students in the same school.

Consider three possible distributions of students across schools. Each of these distributions will generate a different community, which we denote by C_1, C_2, C_3. In Table 4 we illustrate the distribution of social groups across the space for each of them. Each column represents a different community, and then we compute segregation on each of them separately. Each row corresponds to a location in the space (i.e. a school). The triplets on each cell correspond to the share of low, middle and high income students on each school, for a given community. For instance, in the first community, schools 1 and 2 have only low income students, schools 3 to 6 only middle income, and schools 7 and 8 only high income. In community C_2 there is some mixing, so for instance half of the students in schools 1 and 2 have low income, and half have middle income.

Table 4: Distribution of students (As a share of school capacity)

	C_1	C_2	C_3
School 1	(1,0,0)	($\frac{1}{2}$,$\frac{1}{2}$,0)	(1, 0,0)
School 2	(1,0,0)	($\frac{1}{2}$,$\frac{1}{2}$,0)	($\frac{1}{2}$,$\frac{1}{2}$,0)
School 3	(0,1,0)	($\frac{1}{2}$,0,$\frac{1}{2}$)	($\frac{1}{2}$,0,$\frac{1}{2}$)
School 4	(0,1,0)	($\frac{1}{2}$,0,$\frac{1}{2}$)	(0,1,0)
School 5	(0,1,0)	(0,1,0)	(0,1,0)
School 6	(0,1,0)	(0,1,0)	(0,1,0)
School 7	(0,0,1)	(0,$\frac{1}{2}$,$\frac{1}{2}$)	(0,$\frac{1}{2}$,$\frac{1}{2}$)
School 8	(0,0,1)	(0,$\frac{1}{2}$,$\frac{1}{2}$)	(0,0,1)

The distribution of students across schools generates probabilities of observing each type interaction. For instance, the probability of observing an interaction between two low income children in district C_1 is 1/4 (they can meet in two out of eight schools). This is the aggregate intensity function $\tilde{\mu}_\Sigma$ defined in equation 2.2, when μ is correctly normalized. Let $\tilde{\mu}^1_\Sigma$ be the aggregate intensity function of community C_1. This function satisfies $\tilde{\mu}^1_\Sigma(l,l) = \tilde{\mu}^1_\Sigma(h,h) = 1/4$; $\tilde{\mu}^1(m,m) = 1/2$, and $\tilde{\mu}^1_\Sigma(\sigma,\sigma') = 0$ for $\sigma \neq \sigma'$.

Following the same reasoning for the second community, we obtain $\tilde{\mu}^2_\Sigma(l,m) = \tilde{\mu}^2(l,h) = \tilde{\mu}^2_\Sigma(h,m) = \tilde{\mu}^2_\Sigma(m,m) = 1/4$, and $\tilde{\mu}^2_\Sigma(\sigma,\sigma) = 0$ for $\sigma \in \{l,h\}$. It is clear that community 1 is more segregated than community 2: in the first one students with different income levels do not interact, while in the second one there is some mixing.

Now suppose we are interested in combining communities 1 and 2, to obtain a community $C = \alpha C_1 + (1-\alpha)C_2$ with $\alpha = 1/2$. This is analogous to combining the interaction intensities generated by communities C_1 and C_2: the new community C is consistent with a new intensity $\tilde{\mu}$ such that $\tilde{\mu}(\sigma,\sigma') = \alpha\tilde{\mu}^1(\sigma,\sigma') + (1-\alpha)\tilde{\mu}^2(\sigma,\sigma')$, for all $\sigma,\sigma' \in \Sigma$.

Note that this combination generates community C_3, represented in the third column in Table 4. The interaction intensities are as follows:

$$\tilde{\mu}^3(l,l) = \tilde{\mu}^3(l,m) = \tilde{\mu}^3(l,h) = \tilde{\mu}^3(m,h) = \tilde{\mu}^3(h,h) = \frac{1}{8};\ \tilde{\mu}^3(m,m) = \frac{1}{4}. \tag{C.1}$$

C.2 Equivalent Metrics

Lemma 2 *Let $f, g : \mathbb{R} \to \mathbb{R}$ be increasing, continuous and subadditive functions. Then, the order of segregation is preserved for any metrics $d'_\Gamma = f(d_\Gamma)$, $d'_\Lambda = g(d_\Lambda)$.*

Proof. For any function f we can do a first-order approximation around $E(d)$, so that

$$g(d) = g(E(d)) + (d - E(d)) \left.\frac{\partial g(d)}{\partial d}\right|_{E(d)}$$

Thus, without loss of generality fix d_Γ, and take $g(d_\Lambda)$. Then,

$$\begin{aligned} cov(g(d_\Lambda), d_\Gamma) &= cov\left(g(E(d_\Lambda)) + (d - E(d_\Lambda)) \left.\frac{\partial g(d_\Lambda)}{\partial d_\Lambda}\right|_{E(d_\Lambda)}, d_\Gamma\right) \\ &= cov\left(g(E(d_\Lambda)), d_\Gamma\right) + cov\left((d - E(d_\Lambda)) \left.\frac{\partial g(d_\Lambda)}{\partial d_\Lambda}\right|_{E(d_\Lambda)}, d_\Gamma\right) \\ &= \left.\frac{\partial g(d_\Lambda)}{\partial d_\Lambda}\right|_{E(d_\Lambda)} cov\left(d_\Gamma, d_\Gamma\right) \end{aligned}$$

Thus, the order is preserved. Moreover, $g(d_\Gamma)$ and d_Γ are equivalent metrics. Following the same reasoning for d_Λ, we get the result. □

References

Agostini, C., Hojman, D., Román, A. & Valenzuela, L. (2016), 'Segregación residencial de ingresos en el gran santiago, 1992-2002: Una estimación robusta', *EURE* **42**(127), 159–184.

Alesina, A. & La Ferrara, E. (2000), 'Participation in heterogeneous communities', *The Quarterly Journal of Economics* **115**(3), 847–904.

Alesina, A. & La Ferrara, E. (2005), 'Preferences for redistribution in the land of opportunities', *Journal of public Economics* **89**(5), 897–931.

Campante, F. R. & Hojman, D. A. (2013), 'Media and polarization: Evidence from the introduction of broadcast tv in the united states', *Journal of Public Economics* **100**, 79–92.

Coppedge, Michael, e. a. (2020), 'Varieties of democracy (v-dem) project'.

DellaVigna, S. & Kaplan, E. (2007), 'The fox news effect: Media bias and voting', *The Quarterly Journal of Economics* **122**(3), 1187–1234.

Duncan, O. D. & Duncan, B. (1955), 'A methodological analysis of segregation indexes', *American sociological review* **20**(2), 210–217.

Easterly, W. & Levine, R. (1997), 'Africa's growth tragedy: policies and ethnic divisions', *The Quarterly Journal of Economics* **112**(4), 1203–1250.

Echenique, F. & Fryer Jr, R. G. (2007), 'A measure of segregation based on social interactions', *The Quarterly Journal of Economics* **122**(2), 441–485.

Frankel, D. M. & Volij, O. (2011), 'Measuring school segregation', *Journal of Economic Theory* **146**(1), 1–38.

Gentzkow, M. & Shapiro, J. M. (2011), 'Ideological segregation online and offline', *The Quarterly Journal of Economics* **126**(4), 1799–1839.

Hong, L. & Page, S. E. (2001), 'Problem solving by heterogeneous agents', *Journal of Economic Theory* **97**(1), 123–163.

Hong, L. & Page, S. E. (2004), 'Groups of diverse problem solvers can outperform groups of high-ability problem solvers', *Proceedings of the National Academy of Sciences of the United States of America* **101**(46), 16385–16389.

James, D. R. & Taeuber, K. E. (1985), 'Measures of segregation', *Sociological Methodology* **15**, 1–32.

Kennedy, P. J. & Prat, A. (2019), 'Where do people get their news?', *Economic Policy, Volume 34, Issue 97* .

Levy, G. & Razin, R. (2019), 'Echo chambers and their effects on economic and political outcomes', *Annual Review of Economics* **11**, 303–328.

Massey, D. S. & Denton, N. A. (1988), 'The dimensions of residential segregation', *Social Forces* **67**(2), 281–315.

Mullainathan, S. & Shleifer, A. (2005), 'The market for news', *American Economic Review* **95**(4), 1031–1053.

Owens, A., Reardon, S. F. & Jencks, C. (2016), 'Income segregation between schools and school districts', *American Educational Research Journal* **53**(4), 1159–1197.

Putnam, R. D. (2007), ‘E pluribus unum: Diversity and community in the twenty-first century the 2006 johan skytte prize lecture’, *Scandinavian political studies* **30**(2), 137–174.

Reardon, S. F. & Owens, A. (2014), ‘60 years after brown: Trends and consequences of school segregation’, *Annual Review of Sociology, 40:1, 199-218* .

Stroud, N. J. (2008), ‘Media use and political predispositions: Revisiting the concept of selective exposure’, *Political Behavior* **30**(3), 341–366.