ORIGINAL ARTICLE

Does music soothe the soul? Evaluating the impact of a music education programme in Medellin, Colombia

Jonathan Daniel Gómez-Zapata[1,2] · **Luis César Herrero-Prieto**[2,3] · **Beatriz Rodríguez-Prado**[3]

Abstract

Numerous studies have borne out the effects of cultural and music education on individuals' well-being, considering music as a mainly systematic practice or skill or as established educational supply. However, few studies assess the impact of music programmes designed to achieve specific goals, where music is considered as a tool for social change. As a case study, we take the Medellin Music School Network (Colombia), whose education programme for music initiation has been running for 23 years. Our aim is to evaluate the economic and social impact generated by participating in this programme. We use a quasi-experimental propensity score matching technique as the evaluation method. Results show that the programme significantly reduces the probability of participants' becoming involved in conflict, added to which they perceive a better quality of life. Students achieve better academic performance and intensify cultural consumption and participation in artistic activities. Institutional efficacy is reflected through beneficiaries expressing a positive and significant willingness to pay in order to maintain the programme. The work also aims to evidence the usefulness of the methodology for evaluating the impact of cultural policies, particularly in developing areas.

Keywords Cultural economics · Music training · Cultural policy evaluation · Propensity score matching · Efficacy

✉ Jonathan Daniel Gómez-Zapata
jdgomezz@unal.edu.co

Luis César Herrero-Prieto
herrero@emp.uva.es

Beatriz Rodríguez-Prado
bprado@eco.uva.es

1 Research Group in Economics, Culture and Policies, Facultad de Ciencias Humanas y Económicas, National University of Colombia in Medellin (UNAL), Carrera 65 # 59A–110, Bloque 46, Medellín, Colombia

2 Research Group in Cultural Economics, University of Valladolid (UVa), Valladolid, Spain

3 Department of Applied Economics, University of Valladolid (UVa), Valladolid, Spain

JEL Classification Z11 · Z18 · H53 · O12 · C13

1 Introduction

Music is linked to human senses and emotions and is one of the most important manifestations of mankind's creativity as well as being a factor that forges individual identity and realisation (Levinson 2014). Moreover, music also has implications in the social, cultural and economic configuration of given areas, such that it helps to define collective and geographical cultural idiosyncrasy (Mellander et al. 2018), and may also to some degree help to shape long-term economic development for certain areas (Falck et al. 2011, 2018; Backman and Nilson 2018).

As regards the first aspect, numerous studies have shown that music is positively linked to individual achievement. Hallam (2010) draws on clinical research to demonstrate that an active commitment to music has a positive impact on the development of people's cognitive and non-cognitive skills. Measuring the effects of music on education has attracted the attention of various researchers such as Johnson and Memmott (2006) and Southgate and Roscigno (2009) who analyse music participation patterns and find that students who receive music training in schools in the USA perform better in reading, mathematics and languages. Eerola and Eerola (2014) analyse the quality of school life in ten schools in Finland and find that pupils who receive music tuition evidence a direct correlation with the quality indicator, particularly because they obtain greater overall satisfaction with school and are more committed to their studies. Yang (2015) examines the impact of music tuition on educational performance among infants in Germany. Using multivariate regressions and individual fixed effects, the author finds that children involved in music activities (playing an instrument or singing) perform better academically during their teenage years.

However, not only has it been shown that there is a strong link between music education and cognitive development, but also that merely being involved in music is a form of cultural capital accumulation, through experience, which stimulates students' social development (Southgate and Roscigno 2009). Broh (2002) claims that children and teenagers who take part in music courses build a stronger fabric and social support group with colleagues and parents and are less prone to low self-esteem and emotional decontrol. In a similar vein, Rentfrow and Gosling (2003) and Hille and Schupp (2015) agree that engaging in music helps to boost empathy with others, develops various personality aspects and enables those involved to judge both their own and others' achievements better, stimulating competitiveness and ambition. These non-cognitive effects are so important that various studies have shown how music education programmes might prove particularly suited to dealing with risk factors amongst the young, such as helping to reduce crime levels (Daykin et al. 2013; Bernabé-Villodre 2017), fostering peace amongst communities (Ntarangwi 2015) and improving individuals' socio-emotional health and quality of life (Michalos and Kahlke 2008). In Latin America, certain experiences in music development aimed at promoting social progress have proven successful. These include the National System of Orchestras and

Youth Choirs in Venezuela (Cuesta 2011; Alemán et al. 2017), the Iberia Music Education Programme in El Salvador (Bernabé-Villodre 2017) and the subject of our research, the Medellin Music School Network in Colombia.

Moreover, as regards these social and collective effects, as a result of developing the human and cultural capital acquired through artistic practice, it might also be possible to posit certain economic effects in areas, in a threefold sense. Firstly, people with a higher level of education and accumulated cultural capital are more productive and, therefore, tend to display higher levels of employment and greater income (Berry and Glaeser 2005) as well as cultural spending (Ateca-Amestoy 2009). Furthermore, the most skilled human capital is drawn to regions where there is a greater concentration of artistic and cultural goods and services (Pettit 1999; Backman and Nilson 2018), thereby boosting even further the development of these areas (Tubadji 2012; Bucci et al. 2014). Finally, the level of cultural activity and the artistic background are one of today's sources of creativity and entrepreneurship. The cultural and creative sector thus represents an increasingly significant area of production, also evidencing processes of spatial interaction and economies of agglomeration, principally in urban areas (Boix et al. 2015; Coll-Martínez 2019). Artistic and cultural activities therefore play a key role in territorial and urban economic development (Bille and Schulze 2006; Herrero-Prieto 2011), which has led cities over the last few decades to display a greater interest in fostering the creative sector and particularly in increasing the supply of goods and cultural services. In terms of music as an activity sector and accumulated cultural capital resource, Falck et al. (2011, 2018) show that cultural facilities, and more specifically Baroque operas in German cities, attract specialised human capital and highlight differences in the country's regional economic development. Mellander et al. (2018) also argue that Americans' preferences for music reflect and reinforce broader dimensions of political and economic divides in the USA. Music as a consumption good is also sensitive to market analysis for measuring the economic effects generated by its production and demand, as well as the present-day challenges the industry faces (Vid. Cameron 2016). However, exploring these issues in depth lies outside the scope of the present study, whose main concern is to gauge the short- and medium-term impact of a music initiation programme on the social progress of people involved in experiencing music.

In light of the above, this work seeks to evaluate the economic and social impact generated through participation in a music education programme, specifically the Medellin Music School Network (*Red de Escuelas de Música de Medellín*, Spanish acronym—REMM), in Colombia. This is a public policy programme which has been running for 23 years and which constitutes an emblematic example of cultural action aimed at furthering participants' education, social progress and insertion in a context of non-violence in Medellin, a city primarily known for its long period of violence and insecurity in the late twentieth century. We measure the extent to which such aims are achieved. To do this, we employ the quasi-experimental propensity score matching method, which allows us to identify the set of non-enrolled individuals who most resemble treated individuals, based on socio-economic characteristics. These matched non-enrolled individuals form the comparison group, and the social and economic impact of

a music education programme is then calculated as the mean difference in outcomes between these two groups.

Social impact is measured by analysing four conceptual outcome vectors, which are in turn made up of various specific indicators. We first assess the project's social efficacy and the improvement in the quality of life of its participants. We therefore estimate in the first instance the probability of reducing the number of situations involving conflict or violence to which participants or their families might be exposed, in line with the works of Alemán et al. (2017), Bernabé-Villodre (2017) and Daykin et al. (2013), who evidence that music training programmes for children and teenagers immersed in a vulnerable atmosphere can reduce violence and conflict. We also measure the improvement in participants' quality of life through the possibility of changing the place of residence to better socio-economic urban areas, which generally imply greater and better endowments of public services, as well as improved housing conditions. This analytical purpose is in the line of works such as Pettit (1999), who empirically measures the influence of cultural capital on residential mobility, as well as Hand (2018) and Wheatley and Bickerton (2017), who examine people's well-being arising from their engagement in arts and culture.

Secondly, we measure casual effects on educational and occupational efficacy through participation in the music training programme. Various studies have shown that people who take up music (an instrument and/or voice training) evidence greater stimulation and sharpness in several mental faculties (Blood and Zatorre 2001; Hetland 2000; Rabinowitch et al. 2013; Schellenberg 2011; Swaminathana et al. 2017). This then has a positive effect on students' academic performance (Eerola and Eerola 2014; Gardiner et al. 1996; Yang 2015) and provides better job opportunities in the medium term (Miksza and Hime 2015; Corrales-Herrero and Rodríguez-Prado 2016). We measure this dimension through different specific indicators of academic performance (average marks in studies, finishing their studies within the expected time, obtaining study grants) and employability (professional vocation in arts, entry into the labour market and an income which is above the legal monthly salary).

Third, we seek to measure what effects the music programme has on participants' habits and intensity in terms of cultural consumption, taking up the idea put forward in many studies that an education in music helps to increase personal cultural capital and is therefore able to determine the degree and frequency of participation in various cultural activities (Ateca-Amestoy 2009, 2011; Cabane et al. 2016; Wheatley and Bickerton 2017). In our study, we test the impact on cultural consumption through the effects on participation in various artistic and cultural activities (cinema, theatre, dance, museums, concerts, etc.) or alternative uses (sports, videogames).

Finally, we measure institutional efficacy on the project, perceived as loyalty to the REMM and expressed in terms of willingness to pay through monetary support. This valuation may be seen as a measure of the cultural project's legitimacy (Espinal-Monsalve 2013) as well as the desire to maintain it as a local public good (Gómez-Zapata et al. 2018; Herrero-Prieto et al. 2011).

The results obtained contribute towards the academic debate surrounding the link between cultural and music education and the possibilities of personal and social development, whilst also helping to increase the number of empirical applications of

these techniques in developing countries, as is the case in hand. The work is structured in five sections, as follows. Following on from these considerations concerning the theoretical framework and formulation of the research, the second section explains the case study; in the third section, the methodological approach is developed, specifically, the evaluation method using the propensity score matching technique. The results and discussion concerning the impact evaluation of the music initiation project in Medellin are dealt with in the fourth section, whilst the fifth section offers the main conclusions to emerge from the study.

2 The music initiation programme of the Medellin Music School Network (REMM), Colombia

The REMM (Colombia) is a cultural project involving a music initiation programme, which has been running for 23 years and whose main purpose is to support comprehensive human education. It offers children and youngsters in the city of Medellin an education and a focus through music that is geared towards instilling an appreciation of life, fostering non-violent attitudes and changing young people through education, healthy coexistence and social appropriation.[1]

The REMM was founded in 1996 when Medellin was experiencing its most critical period in terms of safety and coexistence, exclusion and governability, triggered by various groups who operated outside the law. This problem led the authorities, in conjunction with the private sector, to seek effective solutions. To achieve this, the decision was made to draw on the Venezuelan National System of Orchestras and Youth Choirs music education model given that, up to the time, it was the only music training programme in the region that had evidenced social effects, since it had enabled the rehabilitation and prevention of criminal behaviour among participants.[2] As a result, the REMM was originally conceived as a means of socially transforming Medellin. Over the years, it gained international recognition by the Inter-American Development Bank (IDB) in 2002–2003 and by the United Nations Organisation for Education, Science and Culture (UNESCO) in 2015.

The network currently boasts 27 music schools[3] spread throughout the city and covering a wide urban area, as can be seen in Fig. 1. Some 80% of these are located in areas with the lowest socio-economic level, and therefore in the most vulnerable districts. Schools are classified in terms of their artistic projection as follows: 13 bowed string instruments, 13 wind and percussion instruments and one section focusing on Colombian instruments. These schools are conceived as music environments for training and cultural participation, as well as places for collective and inclusive education based on musical creation and performance. Each year a registration period is open to

---

[1] For further details, see www.redmusicamedellin.org.

[2] For an assessment reference for this educational project, see the works of Cuesta (2011) and Alemán et al. (2017) who, respectively, apply cost–benefit analysis and experimental techniques (differences in differences) to gauge the social impact of this programme.

[3] A summary of the main characteristics and content of the music education programme may be found in "Appendix 1".
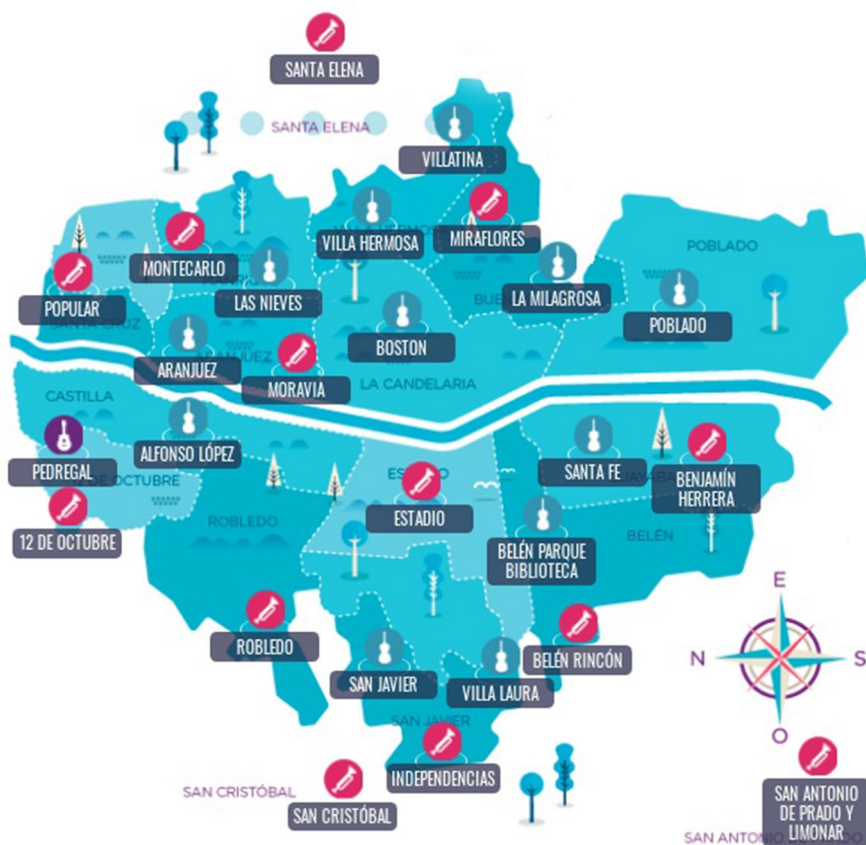
Fig. 1 Location in Medellin of the music schools in the network. *Source*: REMM, www.redmusicamedell
in.org

all children and youngsters aged between 7 and 24, resident in Medellin and enrolled
in a formal education institution. Access to REMM is completely voluntary and free,
and no previous training in music is required: There is therefore no elitist selection
process, nor are there any major differences in terms of quality amongst the various
centres. Through their different channels of information (webpage, social networks,
television, radio, posters, and so on), the Medellin city council and the REMM adver-
tise the registration period and the cultural activities to be held throughout the year.

The REMM's pedagogical model is established in levels, from the least to the
most complex: "seedbed", pre-orchestra or pre-band, orchestra or band and, finally,
integrated ensemble; and involves the children who enrol for the first time taking an
aptitude test to organise them in levels of specific training. Pupils begin the weekly
lessons in groups (of no more than 20) which entail some 4 h, and they are given
instruction in music and instruments, music language and body expression. They
then go on to do courses in musical instrument technique, choral practice and instru-
mental ensemble, in accordance with the particular artistic path they themselves
have freely chosen. The programme embraces a system of merit and recognition. As

a result, students are assessed at each level in terms of their regular attendance[4] and performance of a repertoire chosen by their teachers. The children and youngsters who display the greatest aptitude progress to the next year until they reach one of the 11 integrated ensembles.[5] This requires a minimum of 8 h per week music tuition, yet also affords them the opportunity to meet other students and teachers, but above all offers them the chance to perform at live concerts, thereby enriching the city's cultural and artistic activities. The music training is accompanied by workshops and psychosocial activities, thereby providing students with a comprehensive range of pursuits. This educational process thus seeks to develop skills in the person, in their knowledge and in their actions, and is based on transversal socio-affective interaction as well as the difficulty involved in the repertoire. All of this enables the pupils to enjoy music training and to see themselves as actors from a critical and transformative perspective.

The REMM is an extracurricular programme that is independent from the formal education process. Nevertheless, as being enrolled in formal education is a requirement for accessing and continuing on the programme, it encourages the children and youngsters to remain in their academic programmes; in other words, it discourages them from dropping out of education. Moreover, each participant in the REMM must have a family tutor who is involved in the follow-up meetings and who attends the events and presentation of the results from each school or from the overall programme. In some cases, it is the tutors themselves who make up the music school parents' board and who are most active in following up and supporting their children or relatives.

Since the programme was created, over 15,000 children and youngsters have benefited, with an annual average of some 5000 students enrolled, most of whom belong to the lower socio-economic strata (1, 2 and 3).[6] The annual budget is approximately $ 6000 million pesos (two million dollars), and is used to run and operate the REMM and to organise activities that are freely available to all the public. These include concerts (for example, in addition to each school's own concerts, one of the integrated ensembles offers a concert each month), workshops, festivals, seminars, and academic

---

[4] According to the programme information, the dropout rate is low and stands at less than 10% per year and is generally due to the academic workload at school or university. One reflection of the level of permanence is that 48% of the present-day teaching staff and musical instrument teachers are graduates of the programme.

[5] The 11 integrated ensembles are made up of four symphony orchestras, three choral orchestras and four popular music groups. The presence of popular music groups together with the Colombian music school (Pedregal) enables the REMM to explore autochthonous music, which is reflected in a wide and varied repertoire, ranging from universal classical works to contemporary works composed by school conductors or other musicians in the city, the country and the region.

[6] In accordance with the National Department of Statistics in Colombia, the socio-economic stratification of urban areas classifies the residential dwellings that receive public services into groups. It is mainly carried out in order to levy a differential charge, by different strata, for the public services provided to households, thus enabling subsidies to be allocated and contributions to be levied on these areas. This means that those who have the highest economic capacity pay more for public services and contribute so as to enable the lower strata to make their payments, following a progressive fiscal system. The classification in strata in Colombia is as follows: (1) low–low, (2) low, (3) medium–low, (4) medium, (5) medium–high and (6) high.
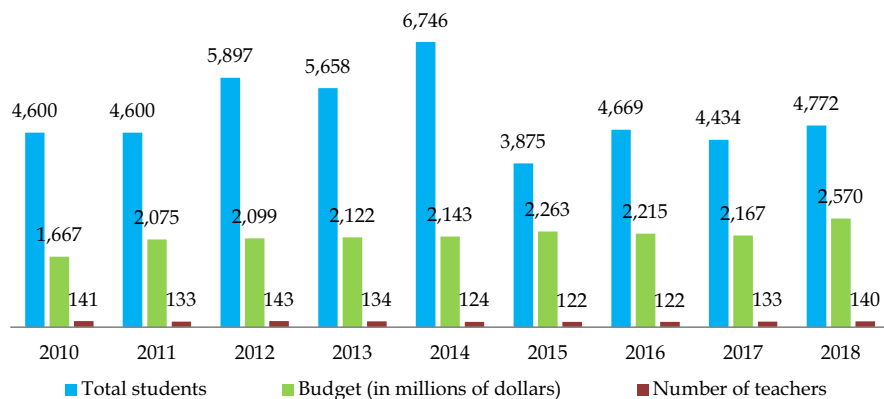
Fig. 2 REMM data. *Source*: Own elaboration based on the management reports of the REMM

conferences as well as national and international tours, where all the members of the various groups travel and perform together, subsequent to obtaining authorisation from their tutors. It helps to create a place where knowledge and experiences can come together, and seeks to enrich the training of students and general public alike.

Figure 2 shows some of the data that are representative of the programme corresponding to its operation in the period 2010–2018.

3 Methodological approach

3.1 Literature review

In order to evaluate the results of the REMM public programme, we use impact evaluation methodologies through which it is possible to measure changes in individuals' well-being that may be attributed to a programme or specific policy. We therefore aim to gauge the causal effect of a programmed cultural action on individuals' results and then as regards the programme's main aims. This provides useful information to those in charge, and is ultimately designed to improve the programme's efficacy (Moral-Arce 2014; Gertler et al. 2017). One of the greatest challenges facing any evaluation is that when estimating a programme's impact, the present situation of the beneficiaries must necessarily be compared to what they would have experienced had they not been beneficiaries of the programme. However, it is impossible to observe individuals in these two scenarios at the same time. Experimental or quasi-experimental evaluation designs must be used to construct a counterfactual scenario that simulates a beneficiary's situation if they had not participated in the programme (Moral-Arce 2014; Gertler et al. 2017). Experimental designs involve a random selection of individuals who are then subject to intervention. It therefore has the advantage of avoiding selection bias at the randomisation level (i.e. before programme execution). Quasi-experimental designs, on the other hand, construct the

counterfactual scenario a posteriori, in other words, after the programme has been implemented. Since participant selection in public programmes is not usually random, the most commonly used evaluation techniques in public policy analysis have been quasi-experimental and have included difference-in-difference (DD) and propensity score matching (PSM) (Smith 2000; Jalan and Ravallion 2003; Pérez-Torres 2015; Gertler et al. 2017). With PSM, each participant is matched with an observationally similar non-participant. The average difference in outcomes after intervention across the two groups is then compared in order to determine the effect of the programme, whereas DD estimates the effect comparing the changes in outcomes over time (before-and-after intervention difference) for the group of participants and a group of non-participants.[7]

As a result, several of these quasi-experimental methods have been applied, particularly as regards culture and education, in order to gauge what effects music education programmes have on development, principally in the area of children and teenagers' cognitive and non-cognitive development. For example, the propensity score matching method is used by Egaña et al. (2013, 2019) to measure the effect on higher education admission test scores of pupils involved in the Curanilahue Youth Orchestra in Chile. Hille and Schupp (2015) also use a variation of this technique to show how children and teenagers who take part in extracurricular music training activities in schools in Germany obtain better marks at school and display more open and ambitious values. Based on panel data, Thomas (2016) applies the difference-in-difference technique to examine what impact the Little Kids Rock programme has on chronic absenteeism in public schools in the USA; and Alemán et al. (2017) also use this technique to show the social effects and the reduction in violence in Venezuela, through participation in the National System of Youth and Child Orchestras Programme.

Nonetheless, a look at the literature shows that, to the best of our knowledge, there are still few cultural project evaluation studies which apply these techniques, particularly in propensity score matching, which is considered one of the most robust and reliable quasi-experimental methods (Jalan and Ravallion 2003). This method constructs the counterfactual scenario by means of a statistical comparison group of non-participants, based on a model of the probability of participating in the treatment, using observed characteristics. The method assumes that non-observed factors do not influence participation in the programme or that at least their influence is collected indirectly by observed variables. By assuming this condition, it is possible to estimate the causal effects on the group of beneficiaries (Rosenbaum and Rubin 1985; Heckman et al. 1998). Participants are therefore matched on the basis of this probability or propensity score to the most similar non-participant. The average treatment effect of the programme is then estimated as the mean difference in outcomes across these two groups.

---

[7] The advantage of PSM is that it only requires information of the outcomes after the programme (cross section data), whereas when applying DD, it is necessary to measure outcomes both before and after the programme (panel data). Moreover, PSM assumes that only observed characteristics can influence participation and outcome, whereas DD can control for unobserved factors. However, for DD to be valid, the comparison group must accurately represent the change in outcomes that would have been experienced by the treatment group in the absence of treatment (i.e. unobserved factors must be time invariant). A more robust method is "matched DD", which combines PSM and DD (Gertler et al. 2017).

Compared to traditional ordinary least square regression, propensity score matching is a more robust method for achieving an unbiased estimate of the treatment effect since it does not assume a functional form of the relation between the outcome and control variables (Caliendo and Kopeining 2008). Moreover, the balance diagnostics for assessing propensity score specification are more transparent than comparable diagnostics for assessing whether an outcome regression model has been correctly formulated. In this regard, goodness-of-fit measures, such as model R2, fail to provide a test of whether the outcome model has been correctly specified. Furthermore, goodness-of-fit tests do not allow us to determine the degree to which the fitted regression model has successfully eliminated systematic differences between participants and non-participants in the REMM. Consequently, with propensity score methods, it is easier to assess whether observed confounding has been adequately eliminated, an aspect which is more difficult to determine when using regression-based approaches (Austin 2011; Gertler et al. 2017).

3.2 Propensity score matching

Empirically, our goal is to estimate the causal impact of participation in the REMM on several dimensions of quality of life and, therefore, its contribution to the social transformation of Medellin. The central question here is how to construct a reliable comparison group. We address this issue by using propensity score matching (Rosenbaum and Rubin 1983; Imbens and Wooldridge 2009; Gertler et al. 2017). The idea is to find, from a group of non-participants, individuals who are observationally similar to participants in terms of characteristics not affected by the programme.

Specifically, in PSM, each participant is matched to a non-participant on the basis of a single propensity score, reflecting the probability of participating, conditional on their different observed characteristics (Rosenbaum and Rubin 1983). To generate propensity scores for our sample, we estimated a probit model in which the dependent variable is a dummy indicating whether the individual participated or not in the REMM (treatment), and the explanatory variables are the socio-economic characteristics of individuals prior to entering the REMM (observed covariates).[8]

Based on this estimated propensity score, we matched each participant in the REMM to the non-participant with the most similar propensity score (one-to-one nearest neighbour, NN(1)). In order to test the sensitivity of the results, we also applied two other matching methods: a variant of one-to-one nearest neighbour that imposes a threshold on the maximum propensity score distance, which in our case is set at 0.01 (Caliper), and Kernel matching, which uses the weighted average of all individuals in the control group to construct the counterfactual outcome of the treated individual. Weights depend on the distance between each individual

---

[8] We also used a logit specification for the propensity score, and the results did not change significantly. These complementary results are available from the authors upon request. Machine learning methods may also provide an alternative nonparametric approach to propensity score estimation, but they are not suitable for small samples (Pirracchio et al. 2015).

from the control group and the participant observation for which the counterfactual is estimated, in such a way that the average places greater weight on people who are close in terms of the propensity score of a treated individual, and lower weight on more distant observations (Caliendo and Kopeining 2008). Although asymptotically the three propensity score matching estimators should yield the same results, in small samples the choice of the matching algorithm may prove to be important since there is a trade-off between bias and variance. Clearly, the first match is always the best and will lead to the least biased estimates, although a many-to-one match will increase sample size and reduce variance (Austin 2011). In our case, as we will see later, the results are qualitatively not sensitive to the specific matching procedure used, such that the choice thereof is unimportant. Once the matched sample is formed, the average treatment effect on those treated is estimated as the mean difference in the outcome variable of the two groups.

There are two critical issues to be considered in any propensity analysis. First, we must ensure there is an overlap in the range of propensity scores across participants in the REMM and non-participants (called "common support") (Rosenbaum and Rubin 1985; Caliendo and Kopeining 2008; Austin 2011), since no inferences about treatment effects can be made for a treated individual for whom there is no individual comparison with a similar propensity score. Therefore, we did not include individuals with a propensity score above the maximum or below the minimum in the potential control group. In our case, there are just a few individuals outside the common support.

Second, the propensity score model must be adequately specified. An adequate propensity model aims to create a matched sample in which the distribution of the observed characteristics is similar between participants and non-participants in the REMM, which is called the balancing property (Rosenbaum and Rubin 1985; Caliendo and Kopeining 2008; Austin 2011). For instance, a propensity should have a similar distribution in the treated and comparison groups. In order to search for such a specification, we followed the strategy proposed by Dehejia and Wahba (1999, 2002). First, we split the sample into blocks of the propensity score so that within each block the mean propensity score was equivalent in the participant and non-participant groups. Second, we tested for equal means of the observed covariates in the two groups within all blocks.

We also applied other methods to check the quality of the matching (post-matching validation). Following Rosenbaum and Rubin (1985), Caliendo and Kopeining (2008), and Austin (2011), we tested for non-significant differences in the mean values of the observed characteristic between the two groups of individuals in the matched sample. We also analysed the reduction in the absolute standardised difference in the mean and median after matching (see Austin 2011 for equations). Standardised differences compare the differences in means (median) in units of the pooled standard deviation (absolute standard deviation). These statistics are not influenced by sample size and allow for the comparison of the relative balance of variables measured in different units. We also considered the reduction in the pseudo-R2 when estimating the propensity score and the lack of joint significance of the explanatory variables when the matched rather than the original sample was used.

Finally, the validity of PSM depends on the conditional independence assumption (also known as "selection-on-observables") (Rosenbaum and Rubin 1985), which states that allocation to treatment is independent of outcomes, conditional on the covariates. This approach only reveals the true causal effect of the programme if all systematic differences in characteristics that influence participation and outcomes can be captured by the observed covariates included. Unfortunately, this assumption is not testable, and we must therefore rely on the observed covariates included in the analysis to cover all the relevant factors that influence treatment and outcomes (Aerts and Schmidt 2008; Gertler et al. 2017). In other words, we must assume there is no selection bias from unobserved characteristics.

In this regard, in the empirical exercise involved in this research we have sought to be thorough when gathering the observed variables, leaving little room for any non-observed factors which might determine participation. Moreover, given the programme's large-scale territorial implementation, inclusion therein is not the result of any exclusive or elitist selection process, but is rather a service that offers free extracurricular training activities on a wide scale. Moreover, as shall also be seen, there are no distinguishing features between the control group and the group of beneficiaries, but that these are in fact fairly homogeneous vis-à-vis characterisation and socio-economic environment variables.

Furthermore, in order to reinforce the validity of our estimates, as will be shown in the results section, we evaluated the degree to which unobserved factors might be biasing PSM results, by applying a combination of PSM and DD methods ("matched DD", Gertler et al. 2017) to an outcome variable measured at the baseline. Results show that potential unobserved factors are negligible.

3.3 Data collection

The empirical information used in this research was obtained based on a survey designed specifically for the case study. The process commenced with a thorough review of the REMM registers in order to select the treatment group, which was made up of pupils first admitted to the programme in 2010 (baseline), and who were at least 14 years of age at that time[9] and who remained in the programme for at least one full academic year (1 year). A total of 150 pupils met these requirements as regards the number of new pupils entering that year, who totalled 288. It was decided to take 2010 as the baseline year for two reasons: firstly, because of the greater quality, confidence and homogeneity of the information provided by the REMM compared to that available for the two previous years, since 2010 witnessed the opening of the last music school (thus making a total of 27), and the University of Antioquia began to run the programme in full, establishing the project of the network curriculum in order to unify the training processes involved and their documentation; secondly, due to the need to use retrospective questions to gather information on certain characteristics of the individuals before their entry into the programme (baseline information). The aim

---

[9] We chose an intermediate age to subsequently check (8 years later) the effects on academic improvement and job opportunities.

was to ensure the remembered baseline year was not too distant in their memory so as to allow those surveyed to locate themselves better in time and thus provide the most recent and trustworthy data possible.[10]

Having identified the registers with the people who met the required previously defined characteristics in 2010, all of them were contacted by telephone and/or email so as to be given the questionnaire. These surveys were applied during the first semester of 2018 (between February and June). The group of non-beneficiaries was constructed using the directed match method (Pérez-Torres 2015); in other words, the beneficiaries surveyed were referenced to between one and three persons (within their family, educational and social context) who had not joined the REMM and who in 2010 fulfilled the same observed characteristics: age, gender, place of residence, education and family environment. These people were also contacted by telephone and/or by email, and the previously referred to questionnaire was applied to them. The final sample obtained contained 180 observations, of which 66 corresponded to REMM participants (44% of the reference population) and 114 to non-beneficiaries (58% of the people surveyed). It would seem that few studies use larger samples to examine the causal effects of music training on participants' cognitive and non-cognitive development (Yang 2015). For example, Gardiner et al. (1996) randomly select 96 first-year music students to estimate their academic performance. Schellenberg (2011) and Swaminathana et al. (2017) assess the correlation between music training and IQ, using samples of 106 and 133 children and youngsters, respectively. Finally, Hetland (2000) conducts a meta-analysis of experimental studies on music performance using samples ranging in size between 12 and 179 pupils. Nevertheless, working with small samples remains a limitation when conducting experimental and quasi-experimental studies.

The questionnaire was divided into two blocks. The first included questions concerning the individuals' situation at the time they joined the REMM (2010, baseline). This section included the observed variables which could theoretically affect both participation in the REMM (treatment) and its possible outcome. These baseline variables included individuals' personal details such as age, sex and level of schooling, and also embraced questions which provided information about their own as well as their family background and social context. Said variables were the educational level of the father and mother, number of persons living in the household and area of residence (commune and socio-economic stratum). In particular, the residence commune in Colombia corresponds to the administrative unit into which the urban area of the cities is subdivided, and which covers certain districts or sectors, which in the case of Medellin is made up of 16 communes.[11] The socio-economic stratum corresponds to the classification of dwellings in six categories, as already mentioned.

---

[10] Notable events were also recalled to help jog the interviewees' memory, such as the IX South American Games or the World Cup held in South Africa that was won by Spain, with the official song sung by the Colombian singer Shakira (*Waka Waka*). The survey model and the primary results thereof are available upon request from the authors.

[11] Commune 1—*Popular*, Commune 2—*Santa Cruz*, Commune 3—*Manrique*, Commune 4—*Aranjuez*, Commune 5—*Castilla*, Commune 6—*Doce de Octubre*, Commune 7—*Robledo*, Commune 8—*Villa Hermosa*, Commune 9—*Buenos Aires*, Commune 10—*La Candelaria*, Commune 11—*Laureles-Estadio*, Commune 12—*La América*, Commune 13—*San Javier*, Commune 14—*Poblado*, Commune 15—*Guayabal* y Commune 16—*Belén*.

The second block in the survey involved questions concerning individuals' situation in 2018, which provided the outcome variables. These variables are grouped into various sections which, as pointed out, address four fundamental analytical vectors through which this cultural programme's economic and social impact was measured: social efficacy, reflecting aspects of improvement in the quality of life, social progress and avoiding reasons for violence on the part of the pupils taking part in the REMM; education and occupational efficacy, which measures educational performance and new medium-term job opportunities for those taking part; cultural consumption effects, related to the frequency and level of participation in various artistic and cultural activities or alternative uses; and, finally, institutional efficiency, perceived as loyalty to the REMM and expressed in terms of willingness to pay through monetary support. Each of these analysis vectors contains specific indicators that have previously been justified (see Sect. 1) and which are now analysed in the form of assessment outcomes.

"Appendix 2" shows the list and content of all the variables employed in the study (socio-economic and programme outcome variables), as well as the descriptive statistics ("Appendix 3") for the sample as a whole and for each analytical group (beneficiaries and non-beneficiaries). It can thus be seen how the full sample ($N = 180$) in the baseline year (2010) shows an equal distribution in terms of gender, with an average age of 16, of whom the majority (80%) were in high school, resided in the urban area of Medellin (73%) and lived in a household made up of an average of four people. In terms of academic achievement, the fathers and mothers had mostly completed high school (27%), followed by elementary school (20%) and, finally, undergraduate (17%). If we look at the level of participants and non-participants in the REMM, we see a similar distribution in almost all of the characteristics. There are perhaps two main differences between these two groups: first, 32% of non-participants lived in rural areas, whereas amongst participants only 18% lived out of Medellin; and second, the level of technological studies of the fathers and mothers and in the group of non-participants is greater (approximately 12%) than that of participants (3%).

4 Analysis and results

The first step towards estimating REMM's economic and social impact is to calculate the likelihood of being allocated treatment, through a probit model.[12] The explanatory variables we included were all the information relative to 2010 (baseline), in other words, prior to joining the REMM. The model finally estimated is shown in Table 1. Even though some of the factors included do not prove to be significant individually, the model as a whole does prove significant and is well specified, as evidenced by the likelihood statistics obtained (LR = 46.70; p value < 0.05). Moreover, it should not be forgotten that the results of the coefficients estimated are accessory and are only used to calculate the likelihood of allocation, which is

---

[12] All of the estimations performed in this study have been carried out using *Stata* software.

Table 1 Probit estimation of propensity score. *Source*: Own elaboration

	Coef.	SE	*z* statistic	*p* value
Gender (*ref.: Female*)				
Male	−0.18	0.23	−0.80	0.42
Age	−0.02	0.06	−0.44	0.66
Educational level (*ref.: Elementary school*)				
Secondary (6th to 9th)	−1.07	0.64	−1.67	0.09*
High school (10th to 11th)	0.08	0.43	0.20	0.84
Family size	0.12	0.08	1.56	0.11
Maximum educational level of the mother (*ref.: Without education*)				
Elementary school (1st to 5th)	−0.05	0.92	−0.06	0.95
Secondary (6th to 9th)	0.85	0.93	0.92	0.36
High school (10th to 11th)	0.15	0.88	0.17	0.86
Technical	−0.16	0.92	−0.18	0.86
Technological	−0.77	0.95	−0.82	0.41
Undergraduate (college)	0.16	0.88	0.19	0.85
Graduate	−0.13	1.01	−0.13	0.89
Maximum educational level of the father (*ref.: Without education*)				
Elementary school (1st to 5th)	0.25	0.60	0.43	0.66
Secondary (6th to 9th)	−1.10	0.73	−1.51	0.13
High school (10th to 11th)	−0.22	0.59	−0.37	0.70
Technical	−0.37	0.65	−0.57	0.57
Technological	−0.76	0.76	−0.99	0.32
Undergraduate (college)	−0.45	0.65	−0.73	0.46
Graduate	−0.42	0.76	−0.55	0.58
Place of residence (*ref.: Commune 5—Castilla*)				
Commune 1—*Popular*	0.44	0.59	0.75	0.45
Commune 3—*Manrique*	0.73	0.53	1.36	0.17
Commune 4—*Aranjuez*	1.17	0.59	1.97	0.04**
Commune 7—*Robledo*	1.22	0.54	2.23	0.02**
Commune 8—*Villa Hermosa*	1.11	0.68	1.64	0.10
Commune 9—*Buenos Aires*	0.72	0.61	1.17	0.24
Commune 10—*La Candelaria*	0.84	0.59	1.42	0.15
Commune 11—*Laureles-Estadio*	0.95	0.61	1.55	0.12
Commune 12—*La América*	0.33	0.66	0.51	0.61
Commune 13—*San Javier*	0.89	0.69	1.28	0.20
Commune 15—*Guayabal*	1.74	0.59	2.96	0.00***
Commune 16—*Belén*	−0.13	0.71	−0.19	0.85
Out of Medellin	0.21	0.43	0.50	0.61
Constant	−0.72	1.49	−0.48	0.62
Number of observations = 180				
LR χ^2 = 46.70**				

The balancing property achieved with five blocks

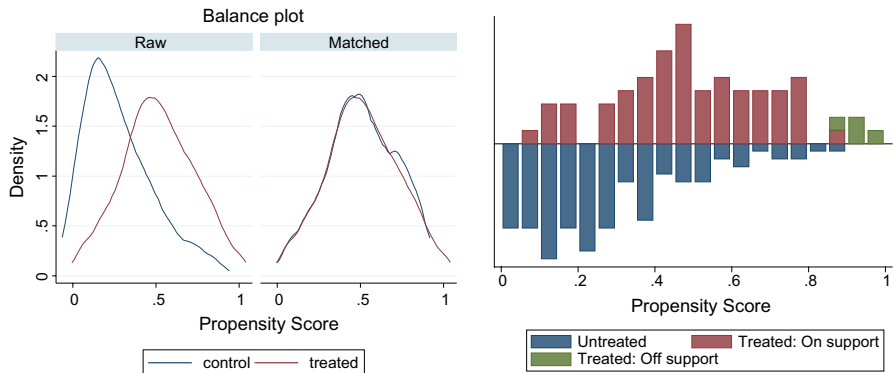*p value < 0.1; **p value < 0.05; ***p value < 0.01

Fig. 3 Region of common support. *Source*: Own elaboration

required to minimise the distance between the two groups, in other words, to perform matching, as previously described.

As also described in the methodology section, several tests were carried out to check the robustness and reliability of the results so as to enable a comparison between participants and non-participants in the REMM before and after matching. The diagnostic statistics on the quality of matching are shown in "Appendixes 4 and 5". "Appendix 4" corresponds to the nearest neighbour criterion (Leuven and Sianesi 2003), and results show that before matching (unmatched sample) most of the variables do not display significant differences at 10%. In fact, this condition is only found in seven of them. This result allows us to conclude that the selection of the sample of non-beneficiaries surveyed through a directed match was suitable for forming two groups (treated and control) that were very similar vis-à-vis their observed features and that; therefore, it is possible to estimate the differences in the results obtained in the two groups.

Secondly, the results also show that once matching has been carried out (matched sample), none of the variables differ significantly in terms of the means of the two groups. In other words, subsequent matching of individuals, based on the nearest neighbour, corrects the remaining difference in the mean of the observed variables between the two groups, thereby reducing selection bias and ensuring the reliability and quality of the results.[13] In this regard, Fig. 3 illustrates the distribution of the propensity score for participants in the REEM (treated) and for non-participants (control) and shows how after matching (matched) a high level of overlap or common support is obtained, estimated in the interval [0.0748, 0.9601]. This implies that only four beneficiaries are outside this region and, therefore, outside the analysis, such that the final operative sample is 176 observations.[14]

---

[13] When using the two other matching techniques posited in the research (Caliper and Kernel), the goodness-of-fit tables yield similar results. These complementary results are available from the authors upon request.

[14] Later on, in the estimations of the causal effects using the nearest neighbour and Kernel techniques, it can be seen that the sample used indeed corresponds to 176 observations ($N=176$), whilst in the case of the nearest neighbour with Caliper (0.01) 13 observations are seen to lie outside the common support region, such that $N=167$.

Finally, as complementary evidence of the inexistence of non-observed relevant factors which might be biasing the results, we applied the matched DD technique, which does take into account non-observed variables (constant over time) for the variables of academic performance.[15] This application (see "Appendix 6") leads to an increase in the probability of improvement in the academic performance for the group of beneficiaries which is similar to that obtained when applying PSM. This shows there are no relevant unobserved factors which, a priori, determine participation in the music education programme and its derivations, or that if there are, they are already indirectly reflected by the observed variables included in the model.

Having checked the quality conditions of matching and of the model, we estimated the mean causal effects on the outcome indicators, in our case, variables related with the four conceptual dimensions of outcomes: social efficacy, educational and occupational efficacy, effects on cultural consumption and institutional efficacy. Shown below are the results to emerge from evaluating the impact of the REMM on these four sections.

4.1 Social efficacy

This section evaluates to what extent the aims of the REMM programme, related to improving coexistence and keeping youngsters away from violence and armed conflicts as well as seeking better quality of life conditions, have been achieved. Table 2 displays the results of the matching procedures through the three techniques applied, evidencing consistency in all cases, and with only minor differences between them. The results reported include the estimated average treatment effect on those treated (ATT) and the estimated mean value of the outcome variable for participants in the REMM (treated) and for the counterfactual situation (control), in other words, the mean value of the outcome variable had they not participated in the REMM. In this way, we provide information on the impact of the programme together with the actual magnitude of the outcome variable analysed.

The results point to two significant effects: first, a reduction in the probability of beneficiaries of REMM and their families being involved in situations of conflict and/or violence. More specifically, if we focus on the nearest neighbour estimations, the reduction in the probability of being involved in situations of conflict and/or violence is on average 19.35 percentage points. This value is obtained from the difference between the probability of being involved in situations of conflict and/or violence between the two groups. As seen in Table 2, said estimated probability for participants in the REEM (treated) is 0.129, whereas the estimated probability had they not participated in the REMM (control) rises significantly, to 0.3225. Put differently, the mean effect of treatment is negative and statistically significant [ATT $= -0.1935$, $p < 0.05$].

---

[15] It should be pointed out that it was not possible to apply this technique to all the outcome vectors, since the variable must be known both before and after treatment. The early age at which pupils join the REMM in the baseline (on average, when aged 14) means that certain variables (related to cultural consumption, WTP, labour market, and so on) lie outside the individual's area of decision at the time, although this not the case for the proxy of academic performance (being enrolled in the school year corresponding to their age).

Table 2 Average treatment effect on the variables related to social efficacy. *Source:* Own elaboration

Outcome variables	NN(1)[a]				NN(1) with Caliper[b]				Kernel[c]			
	N=176				N=167				N=176			
	Treated	Controls	ATT	T stat	Treated	Controls	ATT	T stat	Treated	Controls	ATT	T stat
Being involved in situations of conflict and/or violence	0.1290	0.3225	−0.1935	−2.21**	0.1132	0.3301	−0.2169	−2.34**	0.1290	0.2922	−0.1632	−2.13**
Changing place of residence to areas of a higher stratum	0.2258	0.0080	0.2177	3.73***	0.2264	0.0094	0.2169	3.42***	0.2258	0.0083	0.2174	3.63***
Housing conditions	1.0626	1.0511	0.0114	0.42	1.0537	1.0435	0.0101	0.34	1.0626	1.0563	0.0062	0.26

*p value<0.1; **p value<0.05; ***p value<0.01

[a]Nearest neighbour

[b]Caliper of 0.01

[c]Epanechnikov Kernel with bandwidth of 0.06

This result demonstrates the efficacy of the REMM programme in one of its principal aims, namely to keep children and teenagers away from the city's conflicts and violence. This effect might be explained by the educational model implemented by the REMM and in which music education is complemented by acquiring social and civic values such as tolerance, recognition of others, respect for diversity, teamwork, solidarity and social inclusion, among others values which support the transformation of each individual's immediate (Conejo-Rodríguez 2012). In this way, this public policy's effectiveness may be summed up in one of its most representative slogans: "those who play an instrument will never wield a weapon". This result concurs with the conclusions of the studies by Ntarangwi (2015), Alemán et al. (2017) and Bernabé-Villodre (2017) who evidence that music training programmes for children and teenagers who are in vulnerable situations are able to reduce violence, conflicts and crime.

Second, it was found that beneficiaries of the REMM evidence a 21.77 percentage point increase in the probability of changing their place of residence to districts of a better socio-economic stratum. More specifically, as can be seen in Table 2, the value for that probability is estimated to be 0.2258 for participants and 0.0080 for similar non-participants (control). This points to an improved quality of life for beneficiaries and their families, in other words, a positive impact of taking part in the REMM in terms of quality of life and social progress, as a result of moving to areas of a higher stratum in which there is better access to public services (such as electricity, running water, garbage collection) together with a greater number of public amenities (such as parks, hospitals, schools, shopping centres) compared to where they lived previously, which allows them to enjoy a new and varied range of goods and services. This is in line with the idea put forward by Pettit (1999) who affirms that cultural capital is a consistent predictor of residential mobility patterns, and with the findings of the recent study by Falck et al. (2018) who evidence that people with greater human development look to settle in areas that offer more amenities and, therefore, a wider range of social and cultural goods and services.

Another variable we examine in this area of social progress and quality of life concerns housing conditions. Estimations show that the results reflect no significant treatment effect. Nevertheless, the positive sign may suggest a direct relation: in other words, participating in the REMM may increase the probability of living in dwellings that contain more spaces and are better equipped in terms of domiciliary services, which would enable them to obtain more commodities in the physical infrastructure and so improve their quality of life.

4.2 Educational and labour efficacy

Here we evaluate the effects which taking part in the REMM might have on the cognitive and social development of beneficiaries, measured in terms of academic performance and employability. Table 3 includes the results of the matching procedures for the three techniques applied, which display consistency in all cases, with only minor differences between them. It also shows the mean estimated values for the outcome variables for the treatment and control groups.

Table 3 Average treatment effect on the variables related to educational and labour efficacy. *Source*: Own elaboration

Outcome variables	NN(1)[a]				NN(1) with Caliper[b]				Kernel[c]			
	N=176				N=167				N=176			
	Treated	Controls	ATT	T stat	Treated	Controls	ATT	T stat	Treated	Controls	ATT	T stat
Average marks in their studies	4.2058	4.1477	0.0580	0.98	4.2128	4.1203	0.0924	1.46	4.2058	4.1448	0.0609	1.13
Finishing studies within the expected time	0.7580	0.4596	0.2983	2.99***	0.7169	0.4339	0.2830	2.60***	0.7580	0.5355	0.2225	2.46**
Obtaining study grants	0.2580	0.1209	0.1370	1.82*	0.2264	0.1132	0.1132	1.42	0.2580	0.1395	0.1185	1.55
Professional vocation-studying arts	0.1612	0.0080	0.1532	2.91***	0.1698	0.0094	0.1603	2.76***	0.1612	0.0148	0.1464	2.70***
Entry into the labour market	0.6774	0.5806	0.0967	0.95	0.6981	0.5754	0.1226	1.14	0.6774	0.5841	0.0932	1.01
Income higher than a legal monthly salary	0.4032	0.3387	0.0645	0.64	0.4150	0.3396	0.0754	0.70	0.4032	0.3101	0.0930	1.00

*p value < 0.1; **p value < 0.05; ***p value < 0.01

[a]Nearest neighbour

[b]Caliper of 0.01

[c]Epanechnikov Kernel with bandwidth of 0.06

When focusing on the estimations through the nearest neighbour technique, we find three significant treatment effects. First, there is a 29.83 percentage point increase in the probability of finishing schooling in the stipulated number of years (mean value for treated, 0.7580; for controls, 0.4596). In other words, beneficiaries of the REMM are more likely to finish their academic studies and not to delay the possibility of taking up their chosen profession. Second, there is a 13.70 percentage point increase in the probability of being awarded a grant, such that beneficiaries of the REMM are more likely to be given financial support to help with their professional career training (mean value for treated, 0.2580; for controls, 0.1209). These grants are regularly awarded on merit. It seems that beneficiaries of the REMM tend to get better marks and grades and to evidence better academic performance. This would seem logical when bearing in mind the positive sign found in the effect of the average mark/grade variable [ATT = 0.0580] even though it was not statistically significant. These results concur with the findings reported in other case studies conducted in Chile (Egaña et al. 2013, 2019), Germany (Hille and Schupp 2015; Cabane et al. 2016) and the USA (Chen et al. 2017; Southgate and Roscigno 2009), where a significant causal relationship was shown to exist between participation in music activities and the development of cognitive skills through better academic performance, especially in maths and reading tests.

The third significant effect corresponds to the 15.32 percentage point increase in the probability of studying and taking up a professional career in an area related to music, artistic and creative training. Beneficiaries of the REMM display a greater preference for enrolling on academic programmes related to degrees in music, plastic arts, music IT and sound engineering, among others. Table 3 shows how that probability for beneficiaries is 0.1612 compared to similar non-participants (control) whose probability of enrolling in artistic and musical studies is only 0.0080. A study carried out by Miksza and Hime (2015) indicates that half of these music students continue their professional career in music education. This was also stated by Arias-Pérez (2013) in their study into REMM, where they reported that 48% of all teachers in 2013 are graduates of this programme. In other words, the REMM can have a direct effect on its beneficiaries' career. Moreover, the musical skills acquired during the programme not only enhance their curriculum but also enable them to become involved in a new cultural market in Medellin, foster academic training in the arts and culture and broaden the scope of activities in the music sector, which might prove to be a key source of income.

With regard to the effects on the variables related to joining the labour market and obtaining a minimum income that is equal to a legal monthly salary, no significant differences were found between the two groups. However, the positive sign in the estimated effect does point to a direct relationship between these variables and participation in the REMM. One possible explanation might be that the impact of the REMM has not yet been fully reflected, given that in 2018 the beneficiaries studied had an age equal to or above 22, such that they are likely to have only just finished their university degree and/or that their link with work experience is still recent. As a result, their salary is still likely to be low. It would be interesting to examine this effect in a few years' time. This provides yet further motivation for future impact evaluation studies.

Table 4 Average treatment effect on the variables related to cultural consumption. *Source:* Own elaboration

Outcome variables	NN(1)[a]				NN(1) with Caliper[b]				Kernel[c]			
	N=176				N=167				N=176			
	Treated	Controls	ATT	T stat	Treated	Controls	ATT	T stat	Treated	Controls	ATT	T stat
Doing sports	0.3387	0.6129	−0.2741	−2.66***	0.3396	0.6132	−0.2735	−2.47**	0.3387	0.5692	−0.2305	−2.45**
Attending football matches	2.0483	6.8064	−4.7580	−2.16**	2.2452	8.1886	−5.9433	−2.45**	2.0483	6.1930	−4.1446	−2.04**
Participating in sports competitions	1.2096	1.4274	−0.2177	−0.26	1.2830	1.2547	0.0283	0.03	1.2096	1.6232	−0.4135	−0.59
Cinema attendance	10.8548	7.8467	3.0080	1.24	9.7924	7.0283	2.7641	1.29	10.8548	7.200	3.6543	1.67*
Theatre attendance	4.0161	2.3145	1.7016	1.62	4.2075	1.8301	2.3773	2.02**	4.0161	2.5244	1.4916	1.46
Going dancing	2.1774	1.2096	0.9677	0.84	1.7924	1.0471	0.7452	0.63	2.1774	1.5169	0.6605	0.58
Museum attendance	2.2419	1.8306	0.41129	0.33	2.2264	1.2358	0.9905	0.73	2.2419	2.2675	−0.0256	−0.02
Book reading	43.0483	25.7983	17.25	1.17	48.7547	21.2452	27.5094	1.65*	43.0483	25.4916	17.5567	1.21
Reading newspapers	161.5645	123.8951	37.6693	1.14	166.5471	112.4811	54.0660	1.52	161.5645	123.7440	37.8205	1.32
Attending classical music concerts	5.5322	0.4838	5.0483	2.96***	5.3396	0.6320	4.7075	2.41**	5.5322	0.4099	5.1223	3.03***
Attending concerts	3.2580	1.7741	1.4838	2.26**	3.3018	1.7641	1.5377	2.10**	3.2580	1.7440	1.5140	2.39**
Visiting heritage	1.5161	1.2580	0.2580	0.37	1.5660	0.9339	0.6320	0.82	1.5161	1.8063	−0.2902	−0.12
Videogames	34.4193	26.3145	8.1048	0.52	31.1886	28.8207	2.3679	0.14	34.4193	23.3411	1.10782	0.79

*p value < 0.1; **p value < 0.05; ***p value < 0.01

[a]Nearest neighbour

[b]Caliper of 0.01

[c]Epanechnikov Kernel with bandwidth of 0.06

4.3 Effects on cultural consumption

Table 4 shows the estimations of the mean causal effect of participating in the REMM on cultural and sport consumption in different branches. The results of the three matching procedures applied are included, with the following significant effects being found. First, there is a substitutive relation between participating in the REMM and sport. The results (with the nearest neighbour) show that the probability of a beneficiary doing sport is reduced by 27.41 percentage points, at the same time as their engaging in football-related activities drops by about five times a year. In other words, REMM pupils evidence a different sensitivity, one which is focused on participating in various artistic and cultural activities, with sport emerging as a substitute activity. Although it is true that literature displays a complementary relation between sport and music, since both activities have positive effects on individuals' well-being (Taylor et al. 2015; Cabane et al. 2016; Wheatley and Bickerton 2017), the result we obtain in this study does not discuss this complementarity but merely highlights the rivalry vis-à-vis devoting time between the two. In this regard, beneficiaries of the REMM are willing to sacrifice sport in order to dedicate their time to artistic and cultural activities, as for them this occupation might generate a significant increase in their perception of individual welfare.

Secondly, in this section we also notice an increase in consumption of and loyalty towards cultural goods and services on the part of beneficiaries of the REMM, by attending concerts on average almost twice as often per year as non-participants and five times as often a year in the case of a classical music concert. Positive effects were also found in terms of cinema-going (3.65 times more per year), theatre (2.37 times more per year) and book reading (27.50 times more per year). In other words, participation in artistic and cultural activities is greater if the individual is a beneficiary of the REMM. This result concurs with a number of different studies (Hallam 2010; Sloboda 2012; Ramos-Vargas 2015) which have shown that learning music helps to acquire interiorised cultural capital, whilst also instilling social and civic values in individuals. This cultural capital helps to consolidate a public who demand greater cultural consumption in order to strengthen their cultural trustworthiness, whilst seeking to distinguish themselves from other groups (Bourdieu 2000) and so construct their own cultural capital and reason for loyalty consumption, as has been found in the case of demand for museums (see, Gómez-Zapata et al. 2018 for the case of the Antioquia Museum in Medellin).

4.4 Institutional efficacy

The sense of belonging and loyalty which the REMM triggers amongst its participants is reflected in the willingness to contribute in monetary terms to sustaining the programme which, as shown in Table 5, is $ 53.157 COP (18 USD). This figure is 1.6 higher than is stated by non-beneficiaries. Put differently, if we take the nearest neighbour estimations, the average effect of treatment is positive and statistically significant [ATT = 53,156.66, $p < 0.01$]. In addition, and although not significant, it

Table 5 Average treatment effect on the variables related to institutional efficacy. *Source*: Own elaboration

Outcome variables	NN(1)[a]				NN(1) with Caliper[b]				Kernel[c]			
	N=176				N=167				N=176			
	Treated	Controls	ATT	T stat	Treated	Controls	ATT	T stat	Treated	Controls	ATT	T stat
Willingness to pay	86,394.75	33,238.09	53,156.66	3.04***	88,659.90	28,778.47	59,881.43	2.99***	86,394.75	32,852.21	53,542.53	3.10***
Willingness to support the REMM	0.8548	0.7983	0.0564	0.67	0.8301	0.7547	0.0754	0.81	0.8548	0.7639	0.0908	1.19

*p value <0.1; **p value <0.05; ***p value <0.01

[a]Nearest neighbour

[b]Caliper of 0.01

[c]Epanechnikov Kernel with bandwidth of 0.06

was found that beneficiaries of the REMM evidence a greater probability (approximately 86%) of supporting the programme's continuation in the city.

These results bear out the loyalty of beneficiaries to the REMM and also serve to evidence the programme's legitimacy and social viability since, regardless of whether this cultural project is provided free of charge, this does not mean that it is not valued by its public. Similar results are found in the study carried out by Herrero-Prieto et al. (2011, 2012) who, when estimating the economic value of a classical music festival in *Santiago de Compostela* (Spain), find that 84% of participants expressed a positive willingness to pay for the festival, particularly and with greater certainty on the part of local inhabitants, who therefore felt it to be a local public good, something that might be extensible to our case study.

5 Conclusions

This work provides an evaluation of a cultural policy for an introduction to music and training therein and specifically measures the economic and social impact of the REMM, Colombia, on its beneficiaries under the hypothesis that artistic and music education programmes can have an impact on performance and opportunities for the social progress of those taking part. We take the case of the REMM as emblematic given that it is defined as a tool for social transformation as well as a response to a critical situation in terms of the safety and coexistence, exclusion and governability experienced by Medellin in the late twentieth century.

The review of the literature revealed there to be few studies evaluating the impact of arts and cultural programmes implemented by local government, and even more so in developing countries like Colombia. Consequently, the results to emerge from this work contribute to the academic debate and empirical applications related to these techniques. The choice of the propensity score matching method is felt to be both pertinent and appropriate for the objectives pursued in this research, given that it is grounded on constructing a counterfactual scenario by means of a statistical comparison group (beneficiaries and similar non-beneficiaries) using a participation probability model based on socio-demographic characteristics which are not affected by the programme. This enables us to overcome problems of selection bias present in conventional regression models and thus, through the difference between the mean values of the resulting variables from both groups, pinpoint the most significant mean causal effects of being involved in the REMM. Moreover, the exercise complies with the assumptions of conditional independence and common support required to apply the impact evaluation method and the validation of its outcomes. The use of various matching techniques has allowed us to accredit the robustness thereof.

The distribution of the socio-demographic characteristics in the base year, 2010, both in the group of beneficiaries and non-beneficiaries of the REMM, evidences few differences, thus supporting the suitability of using directed match to construct the control group in this kind of research. In this regard, it should be pointed out that consolidating the baseline information for the two analysis groups has involved one of the greatest efforts, particularly in situations where no previous population data

were available concerning their involvement in the programme or public policy, as is our case. As a result, in order to overcome this limitation, it was necessary to apply a questionnaire with retrospective questions regarding interviewees' initial conditions (2010). In this way, the impact evaluation exercise, perceived in terms of achieving the goals pursued by participation in the REMM, allows us to sum up the results in four aspects: social efficacy and improved quality of life, academic performance and job opportunities, changes in cultural consumption habits and the institutional efficacy of the project.

The social efficacy and improved quality of life may be deduced given that participation in the REMM significantly reduces the possibility of participants' finding themselves involved in conflict and violence, thereby improving levels of well-being and quality of life of the beneficiary population and their families. This might be one of the most valuable results to come out of the study, in that it reflects this cultural policy's significant effectiveness (19.35 percentage points) vis-à-vis complying with one of its principal objectives, which deals with reducing the probability of being involved in situations of conflict or/and violence. In addition, participation in the REMM generates an interest in its beneficiaries to change their place of residence to areas located in sectors with a higher socio-economic level, enabling easier access to social and cultural services and amenities, which ultimately translates to a better quality of life.

Educational and occupational efficacy is reflected by the fact that beneficiaries of the REMM: first, are more likely (29.83 percentage points) to finish their school studies in the stipulated time; second, the probability of being awarded a grant increases by 13.70 percentage points; and, third, there is a 15.32 percentage point increase in the probability of studying and pursuing a professional career in areas related to training in music, the arts and creativity. In other words, the REMM contributes towards the cognitive development of its beneficiaries, making them more efficient at completing their studies and, therefore, helping them to find employment quickly. Moreover, the musical skill acquired during the programme not only enhances their curriculum but also allows them to create and participate in a new cultural market in Medellin, fostering academic training in the arts and culture, expanding activities in the music sector which, among other factors, might prove to be a key source of income. Nevertheless, it should be borne in mind that in this study it was not possible to find a significant mean causal effect in employability and level of income for REMM graduates, this being an issue which provides motivation for future studies. Overall, the results to emerge concerning the REMM's educational and occupational efficacy allow us to argue to the public sector the need to formulate and implement public policies based on music education and training at an early age, given that this enhances people's cognitive abilities, thereby aiding better job placement in the labour market, which can lead to economic growth and a better quality of life.

The effects on the cultural consumption of beneficiaries of this music initiation programme are reflected in the increase in cultural consumption and artistic cluster. That is to say, active participation in the REMM sensitises and intensifies consumption tastes and preferences for artistic and cultural activities, increasing attendance at classical music performances (5.04 times more a year), concerts

(1.48 times more a year), cinema (3.65 times more a year), theatre (2.37 times more a year) and book reading (27.50 times more a year). Through this consumption, beneficiaries feel identified and see themselves as part of a group with the same likes, therefore constituting an important and solid social network. In this regard, they display less interest in creating a loyalty towards other kinds of activities such as sports, the probability of participation in which drops by 27.41 percentage points. Finally, the cultural project's institutional efficacy is reflected by the fact that beneficiaries express a positive and significant willingness to pay of $ 53.157 COP (18 USD) to maintain the programme. This contribution may be seen as a monetary expression of their satisfaction and as a measure of the programme's legitimacy.

As shown previously, the results obtained when measuring the social impact and evaluation of the REMM's efficacy allow us to recognise and value the importance of culture and the arts in countries' social and economic development. In other words, contact between people and communities with artistic and cultural expressions becomes a valuable instrument for the economic development and social transformation of areas. It also helps to justify and legitimise public investment in this type of cultural programme, providing useful information for territorial administrations and cultural managers with regard to maintaining and implementing this type of project. This work thus presents an evaluation of the impact, in terms of goals achieved, of a music education programme whose methodological application may be generalisable to other cultural policies, particularly in the field of developing countries.

Our study is subject to several limitations which in turn offer opportunities for future research. As with any study, our results must be considered in the light of the research sample and data. In this regard, the final sample of participants may not be large, although it covers most of the new students entering the programme in the baseline year of study. Additionally, we were not able to adequately check for any effects on job opportunities, given that graduates are still at an early age. Moreover, in order to provide a more comprehensive evaluation of the programme that goes beyond its impact, future research might focus on a cost–benefit analysis of the REMM, which would require calculating the value of the resources invested in the programme in the medium term. We could also analyse the efficiency of the various schools, using frontier techniques, as a means of evaluating performance in the use of resources in relative terms. Unfortunately, the data requirements for such analyses are still not available. It might even be more pertinent to assess alternative forms of intervention geared towards achieving the same objectives, such as practising sport at an early age, which has also been used as a means of avoiding situations of conflict or violence. It might thus prove interesting to compare whether culture or sport emerges as more effective with regard to social change.

Compliance with ethical standards

Appendix 1: Characteristics of the REMM programme

Mission	Offering children and youngsters in the city of Medellin an education and a focus through music geared towards instilling an appreciation of life, fostering non-violent attitudes and changing young people through education, healthy coexistence and social appropriation
Character	An introduction to music education programme that is extracurricular and independent from the formal education process (school or university)
Publicity	General publicity for registration and activities made by the city council and REMM through information channels such as webpage, social networks, television, radio, posters, and so on.
Intake	Annual, voluntary and free Age range between 7 and 24, resident in Medellin and enrolled in a formal education institution. No previous training in music is required Once registered, students take a music aptitude test to distribute them and form groups of students who have similar levels Each student must present a family member as a tutor to attend and take part in REMM activities
Organisation	27 schools: 13 bowed string instruments, 13 wind and percussion instruments, and one focusing on Colombian instruments Four levels: "seedbed", pre-orchestra or pre-band, orchestra or band, integrated ensembles 11 integrated ensembles: four symphony orchestras, three choral orchestras, and four popular music groups Wide and varied repertoire based on universal classical works and including contemporary works composed by musicians from the city, the country and the region; this also provides for the possibility of performing autochthonous works
Training	Teaching model: collective training programme in an introduction to music and instruments, music language, body expression, instrument technique, choral practice and instrument ensembles; complemented by workshops and psychosocial activities. Students choose their instrument in accordance with the school at which they enrolled Assessment: system of merit and recognition which mainly involves regular attendance and final presentation at the end of the year of a repertoire chosen together with their teachers. The children and youngsters who most stand out progress through the years until they come to form part of the integrated ensembles Intensity: group lessons, two sessions per week of approximately 2 h for the basic levels and 8 h a week for the advanced levels The members of the bands, orchestras and/or ensembles travel and perform together in various local, national and international tours
Costs	All of the costs associated with the running and functioning as well as the REMM's cultural programme are met by the Medellin city council. The activities organised as part of the programme are free for all citizens
Average annual data	5000 students 1000 new students 132 teachers Groups of 20 students 55,000 h of lessons	4000 own instruments 10% drop out rate 4100 families involved Budget of two million US dollars 205 direct jobs	500 concerts 45,000 spectators 108 events around the city each year One international tour

Source: Own elaboration based on REMM

Appendix 2: Measures and definition of variables

Variable		Type	Measure
Treatment	Participation in the REMM	Binary	= 1 If you participated in the REMM program
Baseline variables (matching variables)	Gender	Binary	= 1 Man
	Age	Discrete	Age of the interviewee in 2010
	Educational level		
	Secondary (6th to 9th)	Binary	School year in which the interviewee was in 2010 = 1 Secondary
	High school (10th to 11th)	Binary	School year in which the interviewee was in 2010 = 1 High school
	Family size	Discrete	Number of people living with the interviewee in 2010
	Maximum educational level of the mother		
	Elementary school (1st to 5th)	Binary	= 1 Elementary school
	Secondary (6th to 9th)	Binary	= 1 Secondary
	High school (10th to 11th)	Binary	= 1 High school
	Technical	Binary	= 1 Technical
	Technological	Binary	= 1 Technological
	Undergraduate (college)	Binary	= 1 Undergraduate (college)
	Graduate	Binary	= 1 Graduate
	Maximum educational level of the father		
	Elementary school (1st to 5th)	Binary	= 1 Elementary school
	Secondary (6th to 9th)	Binary	= 1 Secondary
	High school (10th to 11th)	Binary	= 1 High school
	Technical	Binary	= 1 Technical
	Technological	Binary	= 1 Technological
	Undergraduate (college)	Binary	= 1 Undergraduate (college)
	Graduate	Binary	= 1 Graduate
	Place of residence		
	Commune 1—*Popular*	Binary	= 1 *Popular*
	Commune 3—*Manrique*	Binary	= 1 *Manrique*
	Commune 4—*Aranjuez*	Binary	= 1 *Aranjuez*
	Commune 7—*Robledo*	Binary	= 1 *Robledo*
	Commune 8—*Villa Hermosa*	Binary	= 1 *Villa Hernisa*
	Commune 9—*Buenos Aires*	Binary	= 1 *Buenos Aires*
	Commune 10—*La Candelaria*	Binary	= 1 *La Candelaria*
	Commune 11—*Laureles-Estadio*	Binary	= 1 *Laureles-Estadio*
	Commune 12—*La América*	Binary	= 1 *La América*

Variable		Type	Measure
	Commune 13—*San Javier*	Binary	= 1 *San Javier*
	Commune 15—*Guayabal*	Binary	= 1 *Guayabal*
	Commune 16—*Belén*	Binary	= 1 *Belén*
	Out of Medellin	Binary	= 1 Out of Medellin
Outcome variables	Social efficacy		
	Being involved in situations of conflict and/or violence	Binary	Direct question of the questionnaire = 1 "You or any of the members of your family group have been involved in a conflict situation (discussion, civil conflict, altercation, violent events, dispute, fight, among others) in the last 5 years"
	Changing place of residence to areas of a higher stratum	Binary	= 1 If the interviewee changed their place of residence from 2010 to 2018, to an area of greater socio-economic status in the city
	Housing conditions	Continuous	For that outcome we built an indicator which allowed us to sum up the battery of questions relating to the number of spaces (living room, bedrooms, kitchens, bathroom, among others) and utilities (such as running water, electricity, garbage collection) in a house
	Educational and labour efficacy		
	Average marks of the studies	Continuous	Direct question of the questionnaire "What was the average overall grade you graduated with?"
	Finishing studies in the expected time	Binary	= 1 If you finish higher studies than the interviewee agreed, in the time established according to your age
	Obtaining study grants	Binary	Direct question of the questionnaire = 1 If you receive a scholarship for the development of your studies
	Professional vocation-studying arts	Binary	= 1 If the interviewee conducted or performs higher studies related to arts
	Entry into the labour market	Binary	= 1 If the interviewee entered the labour market
	Income above a legal monthly salary	Binary	= 1 If the interviewee receives a monthly salary higher than the regulated minimum

Variable	Type	Measure
Cultural consumption		
Doing sports	Discrete	Direct question of the question-
Attending football matches	Discrete	naire:
Participating in sports competi-tions	Discrete	"Number of times you performed or attended the following activi-ties during the last year"
Cinema attendance	Discrete	
Theatre attendance	Discrete	
Going dancing	Discrete	
Museum attendance	Discrete	
Book reading	Discrete	
Reading newspapers	Discrete	
Attending classical music concerts	Discrete	
Attending concerts	Discrete	
Visiting heritage	Discrete	
Videogames	Discrete	
Institutional efficacy		
Willingness to pay	Continuous	Direct question of the question-naire: "How much would be the volun-tary monetary contribution that I would be willing to give once for a year to support the REMM"
Willingness to support the REMM	Binary	Direct question of the question-naire: =1 If "Do you think it is impor-tant that the community in general be involved with the support of the REMM through a voluntary monetary contribution that allows a greater impact on its activities"

Source: Own elaboration

Appendix 3: Descriptive statistics of the baseline and outcome variables

Sample		Complete				Treated				Controls			
		N = 180				N = 66				N = 114			
Variable		Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max
Treatment	Participation in the REMM	0.36	0.48	0	1				1				
Baseline variables (matching variables)	Gender	0.5	0.50	0	1	0.46	0.50	0	1	0.51	0.50	0	1
	Age	16.43	2.55	13	26	16.21	1.90	14	23	16.56	2.86	13	26
	Educational level												
	Secondary (6th to 9th)	0.07	0.25	0	1	0.03	0.17	0	1	0.09	0.29	0	1
	High school (10th to 11th)	0.12	0.33	0	1	0.12	0.32	0	1	0.13	0.33	0	1
	Family size	4.40	1.41	1	11	4.60	1.49	2	10	4.28	1.35	1	11
	Maximum educational level of the mother												
	Elementary school (1st to 5th)	0.13	0.34	0	1	0.13	0.34	0	1	0.14	0.34	0	1
	Secondary (6th to 9th)	0.12	0.32	0	1	0.15	0.36	0	1	0.10	0.30	0	1
	High school (10th to 11th)	0.27	0.44	0	1	0.33	0.47	0	1	0.24	0.43	0	1
	Technical	0.16	0.36	0	1	0.15	0.36	0	1	0.16	0.37	0	1

Sample	Complete				Treated				Controls			
	N=180				N=66				N=114			
Variable	Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max
Technological	0.08	0.28	0	1	0.03	0.17	0	1	0.12	0.32	0	1
Undergraduate (college)	0.14	0.35	0	1	0.15	0.36	0	1	0.14	0.34	0	1
Graduate	0.05	0.21	0	1	0.03	0.17	0	1	0.06	0.24	0	1
Maximum educational level of the father												
Elementary school (1st to 5th)	0.20	0.40	0	1	0.28	0.45	0	1	0.14	0.35	0	1
Secondary (6th to 9th)	0.09	0.29	0	1	0.04	0.20	0	1	0.12	0.32	0	1
High school (10th to 11th)	0.27	0.44	0	1	0.30	0.46	0	1	0.25	0.43	0	1
Technical	0.10	0.30	0	1	0.10	0.31	0	1	0.10	0.30	0	1
Technological	0.07	0.25	0	1	0.03	0.17	0	1	0.09	0.29	0	1
Undergraduate (college)	0.17	0.38	0	1	0.16	0.37	0	1	0.18	0.38	0	1
Graduate	0.03	0.19	0	1	0.03	0.17	0	1	0.04	0.20	0	1
Place of residence												
Commune 1—Popular	0.05	0.21	0	1	0.06	0.24	0	1	0.04	0.20	0	1
Commune 3—Manrique	0.07	0.26	0	1	0.09	0.28	0	1	0.07	0.25	0	1
Commune 4—Aranjuez	0.04	0.20	0	1	0.07	0.26	0	1	0.02	0.16	0	1
Commune 7—Robledo	0.06	0.24	0	1	0.07	0.26	0	1	0.05	0.22	0	1

Sample	Complete				Treated				Controls			
	N=180				N=66				N=114			
Variable	Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max
Commune 8—Villa Hermosa	0.03	0.18	0	1	0.06	0.24	0	1	0.17	0.13	0	1
Commune 9—Buenos Aires	0.05	0.22	0	1	0.04	0.20	0	1	0.06	0.24	0	1
Commune 10—La Candelaria	0.05	0.22	0	1	0.04	0.20	0	1	0.06	0.24	0	1
Commune 11—Laureles-Estadio	0.05	0.21	0	1	0.06	0.24	0	1	0.04	0.20	0	1
Commune 12—La América	0.04	0.20	0	1	0.03	0.17	0	1	0.05	0.22	0	1
Commune 13—San Javier	0.03	0.18	0	1	0.04	0.20	0	1	0.02	0.16	0	1
Commune 15—Guayabal	0.06	0.25	0	1	0.13	0.34	0	1	0.02	0.16	0	1
Commune 16—Belén	0.04	0.20	0	1	0.03	0.17	0	1	0.05	0.22	0	1
Out of Medellín	0.27	0.44	0	1	0.18	0.38	0	1	0.32	0.47	0	1
Outcome variables												
Social efficacy												
Being involved in situations of conflict and/or violence	0.21	0.41	0	1	0.15	0.36	0	1	0.25	0.43	0	1

Sample	Complete				Treated				Controls			
	N=180				N=66				N=114			
Variable	Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max
Changing place of residence to areas of a higher stratum	0.10	0.30	0	1	0.12	0.32	0	1	0.08	0.28	0	1
Housing conditions	1.07	0.12	0.69	1.57	1.05	0.13	0.80	1.57	1.07	0.12	0.69	1.38
Educational and labour efficacy												
Average marks of the studies	4.15	0.27	3.5	4.8	4.2	0.30	3.5	4.8	4.12	0.25	3.5	4.8
Finishing studies in the expected time	0.62	0.48	0	1	0.77	0.42	0	1	0.53	0.50	0	1
Obtaining study grants	0.18	0.39	0	1	0.25	0.44	0	1	0.14	0.35	0	1
Professional vocation-studying arts	0.07	0.26	0	1	0.15	0.36	0	1	0.03	0.18	0	1
Entry into the labour market	0.64	0.48	0	1	0.68	0.46	0	1	0.62	0.48	0	1
Income above a legal monthly salary	0.35	0.48	0	1	0.39	0.49	0	1	0.33	0.47	0	1
Cultural consumption												
Doing sports	0.41	0.49	0	1	0.33	0.47	0	1	0.45	0.50	0	1
Attending football matches	4.70	11.24	0	92	2.22	4.94	0	25	6.14	13.43	0	92

Sample	Complete					Treated					Controls				
	N=180					N=66					N=114				
Variable	Mean	SD	Min	Max		Mean	SD	Min	Max		Mean	SD	Min	Max	
Participating in sports competitions	1.58	3.94	0	30		1.37	2.69	0	16		1.71	4.51	0	30	
Cinema attendance	8.75	10.88	0	100		10.48	13.96	0	100		7.75	8.52	0	50	
Theatre attendance	2.53	5.06	0	40		3.80	6.93	0	40		1.79	3.39	0	20	
Going dancing	1.58	5.86	0	52		2.13	5.15	0	30		1.26	6.23	0	52	
Museum attendance	2.20	5.69	0	50		2.21	2.70	0	10		2.20	6.86	0	50	
Book reading	33.77	75.50	0	365		46.03	88.37	0	365		26.67	66.32	0	365	
Reading newspapers	131.27	149.80	0	365		162.22	153.26	0	365		113.35	145.44	0	365	
Attending classical music concerts	2.20	8.15	0	100		5.31	12.81	0	100		0.40	1.40	0	12	
Attending concerts	2.16	3.23	0	27		3.31	4.48	0	27		1.49	1.92	0	10	
Visiting heritage	2.98	13.64	0	150		1.54	3.10	0	20		3.81	16.95	0	150	
Videogames	28.32	72.93	0	365		36.87	89.12	0	365		23.37	61.55	0	365	
Institutional efficacy															
Willingness to pay	51,811.07	85,081.35	0	700		82,749.62	126,179.6	0	700		33,899.28	37,760.02	0	200	
Willingness to support the REMM	0.78	0.41	0	1		0.83	0.37	0	1		0.75	0.43	0	1	

Source: Own elaboration

Appendix 4: Diagnostic statistics for one-to-one nearest matching procedure (NN(1))

Variable	Sample	Mean			% Reduct		
		Treated	Control	% Bias[b]	Bias	T test[a]	p value
Gender	Unmatched	0.4697	0.5175	−9.5		−0.62	0.539
	Matched	0.4838	0.4516	6.4	32.6	−0.36	0.722
Age	Unmatched	16.212	16.561	−14.4		−0.88	0.378
	Matched	16.258	15.766	20.2	−40.8	1.38	0.171
Educational level							
Secondary	Unmatched	0.0303	0.0964	−27.3		−1.66	0.099
	Matched	0.0322	0.0249	3.3	87.8	0.27	0.788
High school	Unmatched	0.1212	0.1315	−3.1		−0.20	0.842
	Matched	0.1290	0.1129	4.8	−55.6	0.27	0.785
Family size	Unmatched	4.6061	4.2895	22.2		1.45	0.148
	Matched	4.5968	4.5806	1.1	94.9	0.06	0.953
Maximum educational level of the mother							
Elementary school	Unmatched	0.1363	0.1403	−1.1		−0.07	0.941
	Matched	0.1451	0.0806	18.6	−1518.1	1.13	0.260
Secondary	Unmatched	0.1515	0.1052	13.8		0.91	0.364
	Matched	0.1290	0.1532	−7.2	47.7	−0.38	0.702
High school	Unmatched	0.3333	0.2456	19.3		1.26	0.208
	Matched	0.3387	0.3951	−12.4	35.6	−0.65	0.518
Technical	Unmatched	0.1515	0.1666	−4.1		−0.27	0.791
	Matched	0.1451	0.1129	8.8	−112.9	0.53	0.596
Technological	Unmatched	0.0303	0.1228	−35.2		−2.12	0.036
	Matched	0.0322	0.0564	−9.2	73.8	−0.65	0.517
Undergraduate	Unmatched	0.1515	0.1403	3.1		0.20	0.838
	Matched	0.1612	0.1612	0.0	100.0	0.00	1.000
Graduate	Unmatched	0.0303	0.0614	−14.8		−0.92	0.359
	Matched	0.0322	0.0249	3.8	74.1	0.27	0.788
Maximum educational level of father							
Elementary school	Unmatched	0.2878	0.1912	33.8		2.26	0.025
	Matched	0.2741	0.1935	19.7	41.9	1.06	0.293
Secondary	Unmatched	0.0454	0.1228	−28.0		−1.71	0.088
	Matched	0.0483	0.0806	−11.7	58.3	−0.73	0.469
High school	Unmatched	0.3030	0.2543	10.8		0.70	0.483
	Matched	0.2903	0.3306	−9.0	17.1	−0.48	0.631
Technical	Unmatched	0.1060	0.1052	0.3		0.02	0.987
	Matched	0.1129	0.1693	−18.3	−6979.0	−0.90	0.371
Technological	Unmatched	0.0303	0.0964	−27.3		−1.66	0.099
	Matched	0.0322	0.0161	6.6	75.6	0.58	0.563

Variable	Sample	Mean			% Reduct		
		Treated	Control	% Bias[b]	Bias	T test[a]	p value
Undergraduate	Unmatched	0.1666	0.1842	−4.6		−0.30	0.768
	Matched	0.1774	0.0967	21.1	−359.7	1.30	0.195
Graduate	Unmatched	0.0303	0.0438	−7.1		−0.45	0.652
	Matched	0.3226	0.4839	−8.5	−19.0	−0.45	0.651
Place of residence							
Commune 1—*Popular*	Unmatched	0.0606	0.0438	7.5		0.49	0.622
	Matched	0.0645	0.0887	−10.8	−44.5	−0.50	0.616
Commune 3—*Manrique*	Unmatched	0.09091	0.0701	7.6		0.50	0.619
	Matched	0.0967	0.0645	11.8	−55.6	0.66	0.513
Commune 4—*Aranjuez*	Unmatched	0.0757	0.0263	22.5		1.55	0.122
	Matched	0.0806	0.0645	7.3	67.4	0.34	0.732
Commune 7—*Robledo*	Unmatched	0.0757	0.0526	9.4		0.62	0.535
	Matched	0.0806	0.1048	−9.8	−4.6	−0.46	0.646
Commune 8—*Villa Hermosa*	Unmatched	0.0606	0.0175	22.2		1.55	0.122
	Matched	0.0483	0.0645	−8.3	62.5	−0.39	0.700
Commune 9—*Buenos Aire*s	Unmatched	0.0454	0.0614	−7.1		−0.45	0.655
	Matched	0.0483	0.0564	−3.6	49.4	−0.20	0.842
Commune 10—*La Cande-laria*	Unmatched	0.0454	0.0614	−7.1		−0.45	0.655
	Matched	0.0483	0.0564	−3.6	49.4	−0.20	0.842
Commune 11—*Laureles-Estadio*	Unmatched	0.0606	0.0438	7.5		0.49	0.622
	Matched	0.6452	0.0725	−3.6	51.8	−0.18	0.860
Commune 12—*La América*	Unmatched	0.0303	0.0526	−11.2		−0.70	0.486
	Matched	0.0322	0.0322	0.0	100.0	−0.00	1.000
Commune 13—*San Javier*	Unmatched	0.0454	0.0263	10.2		0.69	0.493
	Matched	0.0483	0.0241	12.9	−26.4	0.72	0.475
Commune 15—*Guayabal*	Unmatched	0.1363	0.0263	40.8		2.90	0.004
	Matched	0.0967	0.0806	6.0	85.3	0.31	0.755
Commune 16—*Belén*	Unmatched	0.0303	0.0563	−11.2		−0.70	0.486
	Matched	0.0322	0.1613	8.1	27.8	0.58	0.563
Out of Medellin	Unmatched	0.1818	0.3245	−33.1		−2.09	0.038
	Matched	0.1935	0.2338	−9.3	71.8	−0.54	0.588

Source: Own elaboration

[a]*t* test for equal means between treated and control samples. *p* values < 0.10 indicate significant mean differences at 10%. *p* values > 0.10 indicate non-significant differences

[b]Standardised percentage bias

Appendix 5: Aggregated diagnostic statistics for matching procedures

	Pseudo R2[a]	LR test[b]	$p>\chi^{2b}$	Mean bias[c]	Median bias[c]	Out of common support[d]
Unmatched sample	0.197	46.70	0.045	14.9	11.0	
Matched sample						
NN(1)	0.075	12.88	0.999	8.9	8.4	4
NN(1) with Caliper	0.107	15.66	0.993	10.2	7.6	13
Kernel	0.042	7.30	1.000	6.8	5.9	4

Source: Own elaboration

[a]Pseudo-R2 of propensity score model

[b]Likelihood ratio test of insignificance of all regressors and *p* value

[c]Mean and median bias are summary indicators of the distribution of the abs(bias)

[d]Treated observation out of common support

Appendix 6: Average treatment effects (matched differences in differences)

Finishing studies within the expected time		Treated	Controls	ATT	*T* stat
NN(1)	$N=176$	-0.04838	-0.3225	0.2741	2.21**
NN(1) with Caliper	$N=167$	-0.0566	-0.3396	0.2830	2.15**
Kernel	$N=176$	-0.0483	-0.2845	0.2361	2.49**

Source: Own elaboration

*p value <0.1; **p value <0.05; ***p value <0.01

References

Aerts, K., & Schmidt, T. (2008). Two for the price of one? Additionality effects of R&D subsidies: A comparison between Flanders and Germany. *Research Policy, 37*(5), 806–822.

Alemán, X., Duryea, S., Guerra, N., McEwan, P., Muñoz, R., Stampini, M., et al. (2017). The effects of musical training on child development: A randomized trial of *El Sistema* in Venezuela. *Prevention Science, 18*(7), 865–878.

Arias-Pérez, J. A. (2013). *Gestión Humana: Actualidad y futuro en la Red de Escuelas de Música de Medellín.* Trabajo para optar el título de Especialista en Gestión del Talento Humano y la Productividad. Medellín: Universidad de Antioquia.

Ateca-Amestoy, V. (2009). El capital humano como determinante del consumo cultural. *Estudios de Economía Aplicada, 27*(1), 87–110.

Ateca-Amestoy, V. (2011). Leisure and subjective well-being. In S. Cameron (Ed.), *Handbook on the economics of leisure* (pp. 52–76). Cheltenham: Edward Elgar.

Austin, P. (2011). An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behavioral Research, 46*(3), 399–424.

Backman, M., & Nilson, P. (2018). The role of cultural heritage in attracting skilled individuals. *Journal of Cultural Economics, 42*(1), 111–138.

Bernabé-Villodre, M. (2017). Music education as a tool to improve socio-emotional and intercultural health within adverse contexts in El Salvador. *Procedia—Social and Behavioral Sciences, 237,* 499–504.

Berry, C., & Glaeser, E. (2005). The divergence of human capital levels across cities. *Papers in Regional Science, 84*(3), 407–444.

Bille, T., & Schulze, G. (2006). Culture in urban and regional development. In V. Ginsburgh & D. Throsby (Eds.), *Handbook of the economics of art and culture* (pp. 1051–1099). Amsterdam: Elsevier.

Blood, A., & Zatorre, R. (2001). Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *PNAS—Proceeding of the National Academy of Sciences, 98*(20), 11818–11823.

Boix, R., Hervás, J. L., & De Miguel, B. (2015). Micro-geographies of creative industries clusters in Europe: From hot spots to assemblages. *Papers in Regional Science, 94*(4), 753–773.

Bourdieu, P. (2000). Las formas del capital. Capital Económico, capital cultural y capital social. In P. Bourdieu (Ed.), *Poder, derecho y clases sociales*. Barcelona: Desclée.

Broh, B. (2002). Linking extracurricular programming to academic achievement: Who benefits and why? *Sociology of Education, 75,* 69–95.

Bucci, A., Sacco, P. L., & Segre, G. (2014). Smart endogenous growth: Cultural capital and the creative use of skills. *International Journal of Manpower, 35*(1/2), 33–55.

Cabane, C., Hille, A., & Lechner, M. (2016). Mozart or Pelé? The effects of adolescents' participation in music and sports. *Labour Economics, 41,* 90–103.

Caliendo, M., & Kopeining, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of Economic Survey, 22*(1), 31–72.

Cameron, S. (2016). Past, present and future: Music economics at the crossroads. *Journal of Cultural Economics, 40*(1), 1–12.

Chen, F., Lui, A., Andrade, H., Valle, C., & Mir, H. (2017). Criteria-referenced formative assessment in the arts. *Educational Assessment, Evaluation and Accountability, 29*(3), 297–314.

Coll-Martínez, E. J. (2019). Creativity and the city: Testing the attenuation of agglomeration economies in Barcelona. *Journal of Cultural Economics*. https://doi.org/10.1007/s10824-019-09340-9.

Conejo-Rodríguez, P. A. (2012). El valor formativo de la música para la educación en valores. DEDiCA. *Revista de Educação e Humanidades, 2,* 263–278.

Corrales-Herrero, H., & Rodríguez-Prado, B. (2016). La incidencia de la formación en la edad adulta en la empleabilidad de los individuos: El caso de España. *El trimestre económico, 1*(329), 149–183.

Cuesta, J. (2011). Music to my ears: The (many) socioeconomic benefits of music training programmes. *Applied Economics Letters, 18,* 915–918.

Daykin, N., de Viggiani, N., Pilkington, P., & Moriarty, Y. (2013). Music making for health, well-being and behaviour change in youth justice settings: A systematic review. *Health Promotion International, 28*(2), 197–210.

Dehejia, R., & Wahba, S. (1999). Causal effects in nonexperimental studies: Reevaluating the evaluation of training programs. *Review of Economics and Statistics, 94,* 1053–1062.

Dehejia, R., & Wahba, S. (2002). Propensity score-matching methods for nonexperimental causal studies. *Review of Economics and Statistics, 84*(1), 151–161.

Eerola, P., & Eerola, T. (2014). Extended music education enhances the quality of school life. *Music Education Research, 16,* 88–104.

Egaña, P., Contreras, D., & Valenzuela, J. (2013). *The effect of artistic activities on the development of cognitive and non-cognitive outcomes in Chile: The case of the Curanilahue Youth Orchestra*. Santiago de Chile: Centro de Investigación Avanzada en Educación, Universidad de Chile.

Egaña, P., Contreras, D., & Valenzuela, J. (2019). The impact of art-education on human capital: An empirical assessment of a youth orchestra. *International Journal of Educational Development, 71,* 102105.

Espinal-Monsalve, N. E. (2013). *La Disponibilidad a Pagar como una medida de la Legitimidad. El caso de la Biblioteca Pública Piloto para América Latina y los Parques Biblioteca de Medellín, Colombia* (Doctoral thesis, Universidad Pablo de Olavide, España).

Falck, O., Fritsch, M., & Heblich, S. (2011). The phantom of the opera: Cultural amenities, human capital, and regional economic growth. *Labour Economics, 18*(6), 755–766.

Falck, O., Fritsch, M., Heblich, S., & Otto, A. (2018). Music in the air: Estimating the social return to cultural amenities. *Journal of Cultural Economics, 42*(3), 365–391.

Gardiner, M., Fox, A., Nowles, F., & Jeffrey, D. (1996). Learning improved by arts training. *Nature, 381*(6580), 284.

Gertler, P. J., Martínez, S., Premand, P., Rawlings, L. B., & Vermeersch, C. (2017). *Impact evaluation in practice* (2nd ed.). Washington, DC: World Bank.

Gómez-Zapata, J. D., Espinal-Monsalve, N. E., & Herrero-Prieto, L. C. (2018). Economic valuation of museums as public club goods: Why build loyalty in cultural heritage consumption? *Journal of Cultural Heritage, 30,* 190–198.

Hallam, S. (2010). The power of music: Its impact on the intellectual, social and personal development of children and young people. *International Journal of Music Education, 28*(3), 269–289.

Hand, C. (2018). Do the arts make you happy? A quantile regression approach. *Journal of Cultural Economics, 42*(2), 271–286.

Heckman, J., Ichimura, H., Smith, J., & Todd, P. (1998). Characterizing selection bias using experimental data. *Econometrica, 66*(5), 1017–1098.

Herrero-Prieto, L. C. (2011). La contribución de la cultura y las artes al desarrollo económico regional. *Investigaciones Regionales, 19,* 177–202.

Herrero-Prieto, L. C., Sanz-Lara, J. A., Bedate-Centeno, A., & del Barrio-Tellado, M. J. (2012). Who pays more for a cultural festival, tourists or locals? A certainty analysis of a contingent valuation application. *International Journal of Tourism Research, 14*(5), 495–512.

Herrero-Prieto, L. C., Sanz-Lara, J. A., & Devesa-Fernández, M. (2011). Measuring the economic value and social viability of a cultural festival as a tourism prototype. *Tourism Economics, 17*(3), 639–653.

Hetland, L. (2000). Listening to music enhances spatial-temporal reasoning: Evidence for the 'Mozart Effect'. *Journal of Aesthetic Education, 34,* 105–148.

Hille, A., & Schupp, J. (2015). How learning a musical instrument affects the development of skills. *Economics of Education Review, 44,* 56–82.

Imbens, G., & Wooldridge, J. (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature, 47*(1), 5–86.

Jalan, J., & Ravallion, M. (2003). Estimating the benefit incidence of an antipoverty program by propensity-score matching. *Journal of Business & Economic Statistics, 21*(1), 19–30.

Johnson, C. M., & Memmott, J. E. (2006). Examination of relationships between participation in school music programs of differing quality and standardized test results. *Journal of Research in Music Education, 54*(4), 293–307.

Leuven, E. & Sianesi, B. (2003). psmatch2 version 4.0.12: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing. Download from http://ideas.repec.org/c/boc/bocode/s432001.html. Accessed January 30, 2016.

Levinson, J. (2014). Values of music. In V. Ginsburgh & D. Throsby (Eds.), *Handbook of the economics of art and culture* (Vol. 2). Elsevier: North Holland.

Mellander, C., Florida, R., Rentfrow, P., & Potter, J. (2018). The geography of music preferences. *Journal of Cultural Economics, 42*(4), 593–618.

Michalos, A., & Kahlke, P. (2008). Impact of arts related activities on the perceived quality of life. *Social Indicators Research, 89*(2), 193–258.

Miksza, P., & Hime, L. (2015). Undergraduate music program alumni's career path, retrospective institutional satisfaction, and financial status. *Arts Education Policy Review, 116*(4), 1–13.

Moral-Arce, I. (2014). Elección del método de evaluación cuantitativa de una política pública Buenas prácticas en América Latina y la Unión Europea. *Colección Documentos de Trabajo* nº 6. Madrid: Programa EUROsociAL.

Ntarangwi, M. (2015). Youth, music, and peace building. In J. Wright (Ed.), *International encyclopedia of the social & behavioral sciences* (2nd ed, pp. 842–848). Amsterdam: Elsevier.

Pérez-Torres, F. J. (2015). Evaluación de impacto de política pública del Programa Especial Desarraigados: una aplicación cuantitativa. *Equidad & Desarrollo, 23,* 77–125.

Pettit, B. (1999). Cultural capital and residential mobility: A model of impersistence in place. *Poetics, 26*(3), 177–199.

Pirracchio, R., Petersen, M. L., & van der Laan, M. (2015). Improving propensity score estimators' robustness to model misspecification using super learner. *American Journal of Epidemiology, 181*(2), 108–119.

Rabinowitch, T., Cross, I., & Burnard, P. (2013). Long-term musical group interaction has a positive influence on empathy in children. *Psychology of Music, 41*(4), 484–498.

Ramos-Vargas, M. (2015). Music as a resource to develop cognition. *Procedia—Social and Behavioral Sciences, 174,* 2989–2994.

Red de Escuelas de Música de Medellín (REMM) (2010). Retrieved from www.redmusicamedellin.org. Accessed 8 Feb 2019.

Rentfrow, P., & Gosling, S. (2003). The do re mi's of everyday life: The structure and personality correlates of music preferences. *Journal of Personality and Social Psychology, 84*(6), 1236–1256.

Rosenbaum, P., & Rubin, D. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika, 70*(1), 41–55.

Rosenbaum, P., & Rubin, D. (1985). Constructing a control group using multivariate matched sample methods that incorporate the propensity score. *American Statistician, 39*(1), 33–38.

Schellenberg, E. G. (2011). Examining the association between music lessons and intelligence. *British Journal of Psychology, 102*(3), 283–302.

Sloboda, J. (2012). *La mente musical. La psicología cognitiva de la música*. Madrid: Machado Nuevo Aprendizaje.

Smith, J. (2000). A critical survey of empirical methods for evaluating active labor market policies. *Swiss Journal for Economics and Statistics, 136*(3), 1–22.

Southgate, D., & Roscigno, V. (2009). The impact of music on childhood and adolescent achievement. *Social Science Quarterly, 90*(1), 4–21.

Swaminathana, S., Schellenberg, E. G., & Khalil, S. (2017). Revisiting the association between music lessons and intelligence: Training effects or music aptitude? *Intelligence, 62,* 119–124.

Taylor, P., Davies, L., Wells, D., Gilbertson, J. & Tayleur, W. (2015). *CASE: A review of the social impacts of culture and sport*. Department for Culture, Media and Sport. Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/416279/A_review_of_the_Social_Impacts_of_Culture_and_Sport.pdf. Accessed 1 Apr 2019.

Thomas, M. (2016). Music education and its causal impact on student engagement and success. A program evaluation of little kids Rock. Working paper. Retrieved from https://www.arts.gov/sites/default/files/Research-Art-Works-MississippiState.pdf. Accessed 1 Apr 2019.

Tubadji, A. (2012). Culture-based development: Culture as an encompassing economic factor–empirical evidence for Germany. *International Journal of Social Economics, 39*(9), 690–703.

Wheatley, D., & Bickerton, C. (2017). Subjective well-being and engagement in arts, culture and sport. *Journal of Cultural Economics, 41*(1), 23–45.

Yang, P. (2015). The impact of music on educational attainment. *Journal of Cultural Economics, 39*(4), 369–396.