IZA Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 10929

Segregation and Fertility: The Case of the Roma in Serbia

Marianna Battaglia
Bastien Chabe-Ferret
Lara Lebedinski

JULY 2017

DISCUSSION PAPER SERIES

IZA DP No. 10929

Segregation and Fertility: The Case of the Roma in Serbia

Marianna Battaglia
University of Alicante

Bastien Chabe-Ferret
IRES and IZA

Lara Lebedinski
FREN and IEN

JULY 2017

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Segregation and Fertility: The Case of the Roma in Serbia*

We study the effect of residential segregation on fertility for the socially excluded and marginalized Roma ethnic minority. Using original survey data we collected in Serbia, we investigate whether fertility differs between ethnically homogeneous and mixed neighborhoods. Our results show that Roma in less segregated areas tend to have significantly fewer children (around 0.9). Most of the difference arises from Roma in less segregated areas waiting substantially more after having a boy than their counterparts in more segregated areas. We account for the endogeneity of the level of segregation using (il)legal possibility to build in the area at the time of its creation as an instrument. We find that the true gap due to segregation is actually larger than that estimated by OLS (around 1.4). We finally provide evidence that exposure to the Serbian majority culture is the main mechanism at play, as opposed to differences in opportunity cost of time, migration patterns, family arrangements and returns to education.

JEL Classification: J13, J15, R23, Z10

Keywords: fertility, residential segregation, ethnic minority, culture

Corresponding author:
Marianna Battaglia
Department of Economics (FAE - Fundamentos del Analisis Economico)
University of Alicante
Campus de San Vicente
03080 Alicante
Spain

E-mail: mbattaglia@ua.es

* We are especially indebted to the personnel of the Ipssos Strategic Marketing Agency for the successful data collection. We are grateful to David de la Croix, Anna Sanz-de-Galdeano, Lola Collado, Coralio Ballester, Pedro Albarran and Iñigo Iturbe-Ormaetxe for helpful comments and suggestions. We benefited from the valuable comments of seminar participants at IRES macro lunch seminars, Maastricht University Graduate School of Governance seminars, SAEe 2016 and at the summer accademy on Family in Transition: Context, Values, and Choice (IOS). Financial support from the Spanish MEC (Ref. ECO2014-58434-P) and the Serbian Ministry of education, science and technological development (project number: OI 179015) is gratefully acknowledged. All opinions expressed are of the authors, all errors are our own.

1 Introduction

The Roma population, like many other marginalized minority groups, is characterized by high levels of fertility and severe residential segregation. Yet, both the direction of causality and the mechanism responsible for this correlation remain unclear. Is it that minorities more inclined towards a higher fertility tend to crowd out other types from their neighbourhoods and become segregated as a result or that groups that are initially isolated tend to be biased towards larger families? Alternatively, it could be that a third factor causes both segregation and high fertility.[1]

The aim of this paper is to use variations in the severity of residential segregation of Roma settlements to analyze whether segregation can be responsible for the observed fertility levels.[2] Another goal is to pin down through which mechanisms segregation may operate. Providing answers to these questions is of primary importance to understand whether policies favoring social diversity may be useful to reduce what some consider as a fertility burden, which prevents parents from investing in the quality of their children.[3] Understanding through which mechanism segregation operates is also crucial as policies favoring social diversity may target access to housing, to schools or to jobs and some may prove more efficient than others.

In particular, we have in mind five mechanisms through which segregation may affect fertility: (i) people in less segregated areas may have access to better employment possibilities and therefore have a higher opportunity cost of time (Doepke, 2015); (ii) people in less segregated areas may face higher returns to education and therefore prefer to invest in quality rather than quantity of children (Galor, 2012); (iii) people in segregated areas may be closer from the grandparents' location and raising children would consequently be less costly; (iv) the cost of space could be lower in segregated areas as fewer people desire to live there (Boustan and Margo, 2013; Boustan, 2012); (v) people in less segregated areas may be more exposed to the Serbian majority culture and its low fertility norm.

For the purpose of our analysis, we use primary data collected through an extensive survey conducted in Belgrade, Serbia. In the Fall of 2010, we interviewed 300 Roma households in 13 different settlements of the city. These households were randomly selected among households

[1] Minorities speaking a language that is very distant from that of the majority for instance should tend to be more isolated and disadvantaged on the labor market, which decreases the opportunity cost of having children.

[2] Settlements and neighbourhoods are used as synonyms.

[3] There exists an important literature showing that the decline in fertility known as the demographic transition is a prior to the economic take-off. See for instance the seminal contributions by Galor and Weil (1999, 2000) and subsequent articles building on the issue like Kalemli-Ozcan (2002); Li and Zhang (2007); Klemp and Weisdorf (2012); Cervellati and Sunde (2015).

with at least one child attending primary schools involved in a remedial education program introduced in Serbia in 2009.[4]

We first document that there exists an heterogeneity within the Roma community in Serbia in terms of both segregation and fertility. Residential segregation is measured as the proportion of Roma living in a settlement and we distinguish between *only Roma*, *mostly* or *few Roma*. In our sample, 27% of the households live in *only Roma* neighborhoods, 62% in *mostly Roma* and 9% in *few Roma* settlements. Following a naive hypothesis that segregation is exogenous, we establish that households in *few Roma* settlements tend to have fewer children than those in *mostly Roma* (0.9 of a child). Adding proxies capturing the different mechanisms previously described can only partially account for the gap, but points at the importance of different levels of exposure to the Serbian majority culture. We then investigate whether this fertility gap is accompanied by a preference towards sons rather than daughters. We use a proportional hazard duration model and show that a large part of the gap is due to households waiting significantly longer after having a boy in the *few Roma* settlements.

In a second step, we acknowledge the potential endogeneity of residential segregation and implement an instrumental variable approach. Indeed, living in one settlement or another is not random: residential location is a choice made by the households and the decision of living in a more or less segregated settlement is very likely related to unobserved attributes affecting fertility decisions and gender preferences. There is a potential problem of self-selection into settlements whose direction is a priori unclear. The instrument we use for current segregation is the (il)legal possibility to build in the area at the time of the creation of the settlement. Settling in a place where, on average 50 years ago, authorities had either set a regulation allowing to build, not set any such regulation or instead explicitly forbid it, is strongly correlated with the current residential segregation. A settlement where it was illegal to build at the time of its creation was most likely to become *only Roma* than a settlement created spontaneously with or without some municipality permission and it is still more likely to be an *only Roma* settlement nowadays. However, the (il)legal possibility to build in the area at the time of the creation of the settlement should be orthogonal to current individual characteristics relevant for fertility. Therefore the only channel through which (il)legal possibility to build influences

[4]Data were collected to examine the impact on parental expectations of a remedial education program for primary school-age children targeting the Roma minority group. More detailed can be found in Battaglia and Lebedinski (2015). Our study de facto focuses on the intensive margin of fertility (or fertility of mothers). We discuss this point in Subsection 2.2.

fertility is through the current level of segregation.

We confirm that fertility in *only* and *mostly Roma* settlements is higher than in *few Roma* settlements and that the gap is larger than using straight OLS (around 1.4 children). We therefore conclude that there exists an attenuation bias in our OLS estimates due to self-selection of low fertility individuals into more segregated areas. We then again add successively controls capturing the different mechanisms that could be at work and find even stronger evidence that exposure to Serbian culture is the most important mechanism at play.

Our paper contributes to the developing literature that aims at quantifying the relative importance of cultural norms and economic incentives to explain behavior. de la Croix et al. (2016) build a model of fertility and investment in education that tries to match fertility and schooling data for XIXth century France. They find that the economic model may explain 38% of the cross-county variation in fertility but more than 75% of the variation in schooling decisions. They then correlate the residuals from the model to cultural proxies and find that family structure and cultural barriers are important determinants of fertility. Although pretty recent in the economics literature, the debate is long-standing in demography and sociology.

The sociological literature refers to three hypothesis that may explain fertility differences across ethnic or religious groups: (i) the characteristics hypothesis, which states that, once accounted for differences in socio-demographic characteristics, fertility differences should disappear, (ii) the cultural hypothesis, according to which fertility differences persist due to the slow process of acculturation of minorities to the majority culture and (iii) the minority status hypothesis, which posits that minority group membership may have an independent effect on fertility, either positive or negative due to the desire and perceived possibility of upward social mobility.

These hypothesis have been revisited by the economics literature. The seminal contributions by Fernández and Fogli (2006, 2009) bring to light the correlation between the behavior of second generation migrant women to the US and the Total Fertility Rate in their country of origin to give strong empirical support to the existence of the cultural channel.[5] Chabé-Ferret (2013) show that the characteristics hypothesis does not allow to explain fully the fertility gap between non-Hispanic Whites and African Americans in the US and give some evidence pointing at the importance of the cultural channel. Goldscheider and Uhlenberg (1969) were the first to propose

[5] See Blau et al. (2013); Stichnoth and Yeter (2016); Chabé-Ferret (2016) for further explorations of the cultural channel at play in fertility behavior.

the concept of minority status hypothesis, which gave rise to a substantial literature that tried to prove or disprove them. Chabé-Ferret and Melindi Ghidi (2013) build an economic theory that proposes a mechanism that could underly the minority status hypothesis and find that it is consistent with fertility differentials between the main ethnic groups in the US: non-Hispanic Whites, African Americans, Hispanics and Asians.

In demography, the debate has rather been put in terms of diffusionist versus structuralist views. The former think that fertility obeys to norms that are transmitted across time and space, while the latter believe that fertility responds mainly to economic determinants. It was the purpose of the so-called Princeton study in 1986 to determine whether the demographic transition is Europe was consistent with one or the other view. The study analyzed the timing of the transition across European countries and concluded in favor of the diffusionist view. Our paper in this respect goes in the same direction by giving supportive evidence that exposure to different fertility norms (or lack thereof) may be responsible for the persistently higher fertility of minority groups and in particular of Roma populations.

Our work is also related to the literature on residential segregation and neighborhood effects that studies the relevance of neighborhoods and one's peers in influencing socioeconomic outcomes.[6] For instance, segregation of the African Americans has been identified as one of the reasons for the persistence of inner city poverty in the US (Cutler and Glaeser, 1997). Moreover, the neighborhood where one lives can clearly affect one's labor market (Clark and Drinkwater, 2002; Edin et al., 2003; Bayer et al., 2008; Boeri et al., 2015) and educational outcomes (Card and Rothstein, 2007). The ethnic composition of a municipality can also be important for the quality of local public goods such as schools (Alesina et al., 1999; La Ferrara and Mele, 2006).

Manley et al. (2011) suggests that the evidence base for social mixing is far from robust. Our setting allows to better isolate the effect of segregation on fertility for a minority group, given that we can observe different levels of segregation for the same ethnic minority, which is the largest in Europe. We also contribute to the existing literature by providing primary data in a context where data are scarce. We collected data at a very detailed level of geographical disaggregation - the street - that can be merged with other sources regarding the urban plan at the settlement level in order to get precise information on the context we are studying.

The rest of the paper is organized as follows. Section 2 describes the way the survey has

[6] For an excellent review of the literature on neighborhood effects see Durlauf (2004); Blume and Durlauf (2006).

been designed and the data collected. It provides some descriptive statistics. Section 3 presents the estimation strategy and the results. Section 4 discusses findings and concludes.

2 Data and Descriptive Statistics

2.1 Official data on Roma

Roma are the largest ethnic minority in Europe. In all the countries where they live, they experience severe social exclusion and poverty. They mainly perform low skilled jobs, live in segregated areas of the main cities and do not participate in the political and cultural life (Open Society Institute, 2007). Their living conditions are often so different from those of the majority population that it is difficult to find official data documenting their situation. In the context we are studying, as it is the case for most Central and Eastern European countries where the majority of the Roma population lives, official data on them are scarce and inaccurate. The 2011 Serbian Census counts 147,604 Roma, corresponding to 2.05% of the total Serbian population, while the Open Society Institute (2007) estimates a number between 350,000 and 500,000, approximately 6% of the overall population. In Belgrade, the 2011 Census records 27,325 Roma (1.65% of the population) and the Open Society Institute (2007)'s numbers are three times higher: they are roughly 80,000 (5%).

The UNICEF Multiple Indicator Cluster Survey (MICS) from 2010 and the Living Standard Measurement Survey (LSMS) from 2007 are valuable sources of information on the living conditions of the Roma population in Serbia. However, the MICS does not report information on where the households interviewed live and therefore cannot be used in our study, and the LSMS does not include a representative sample of Roma since those interviewed are only a boosted sample of internally displaced people. We use those sources in order to show the different characteristics of the Roma population with respect to the Non-Roma population. As reported in Table A in the appendix, the average Roma household is composed of 5.6 members versus a national average of 3.5. The average number of children aged 18 or below is 2.4 per Roma households, while the population average is only 0.86. Almost half of the Roma population (43%) is below 18 years old and the average age is 25, whereas the national average is 35. Half of the Roma households are poor: their average consumption is below the absolute poverty line.[7] While, male employment rates are comparable to those of the majority population (56%), yet,

[7] The percentage of the extremely poor among the Roma interviewed in LSMS is 11.9%. Those who are considered extremely poor are those who cannot satisfy even their basic needs for food.

female employment remains very low with only one woman out of ten working versus a national average of 40%. Only 89% of children from Roma settlements aged 6 to 15 attend school and among the adults, 29% have not finished primary school.[8] Conversely, 99% of Non-Roma aged 6 to 15 are enrolled in school and only 4% of adults have not completed primary school.

2.2 The Sample

We use first-hand collected data obtained through a survey conducted with 300 Roma households in 5 municipalities of Belgrade.[9] Our sample is constructed in such a way that all households have at least one child in the lower four grades of primary school in the year of the survey. They were randomly selected among pupils attending primary schools involved in a remedial education program introduced in Serbia in 2009.[10]

The design of the sample implies that all households observed count at least one child. This may distort the representativity of the results in terms of fertility of the whole Roma Population as it de facto removes the extensive margin of the fertility decision (or childlessness as put by Gobbi (2013); Baudin et al. (2015)). Using Census data for 2011, we obtain that childlessness rates of women over 39 in the Belgrade region were of 7% for Roma and 13% for Serbs. Though substantial, the magnitude of this difference is not likely to explain fully the much higher fertility of Roma women. We thus consider that studying only the intensive margin of the fertility decision is an important step per se.

Figure 1 displays a map of Belgrade with the 13 settlements where the survey was carried out. The number of households selected from each settlement is proportional to the size of the settlement. We classify settlements as composed of either *only Roma* people, *mostly Roma* people or *few Roma* people. We have both "official data" from a database of Roma settlements (Society for Improvement of Roma settlements, 2002) and self-reported information. In the survey, we asked respondents whether in their community/neighborhood there were only Roma people or both Roma and Non-Roma, and in the latter case whether Roma were a minority or

[8]In Serbia, school is compulsory until the age of 15 and primary school lasts 8 years. Children enroll at primary school if they are above 6.5 years of age at the start of the scholastic year in September. Since 2010 the attendance of at least 9 months of a free preschool program is compulsory.

[9]The five municipalities are Voždovac, Zvezdara, Zemun, Palilula, and Čukarica. In Belgrade, the city municipalities settlements with more than 15 households or 100 Roma people are 8, 5 of which are in our sample (Jaksic, B. and G. Basic, 2010). The survey took place in Fall 2010. The response rate is 93.46%: 321 households have been contacted and 300 answered. Households were not compensated for their participation.

[10]The Roma Teaching Assistant Program is the main program in Central and Eastern Europe aimed at improving inclusion of Roma in education. For a more extensive description of the program see Battaglia and Lebedinski (2015).

a majority.[11] We use both information and the exact address of their dwellings.

Our data include household members' demographic characteristics, such as education level, religion, language spoken at home and information on their dwellings. We also have detailed information on their settlement of residence from the database of Roma settlements. Besides ethnic composition, the database of Roma settlements provides information on the number of inhabitants and main current utilities. It also reports the year and the way of creation of the settlement, the legal possibility to build and the main purpose of the area at the time of its creation according to the general urban plan.

Panel A in Table 1 reports households' characteristics for the 13 settlements in our sample in column (1). Overall, they are in line with official data (MICS 2010 and LSMS 2007), as reported in Table B in the appendix. On average, women in our sample are 32.5 years old and have spent 5 years at school, likely completing only the first cycle of primary school.[12] Their wealth is measured by an index that ranges from -3.135 to 2.865 based on the presence in the household of various durables and utilities.[13] Households mainly receive income from labor, either in the formal or informal sector, rather than social transfers.[14] Almost all households in the sample are nuclear with on average a little over two adults, although there are a few exceptions with more than four. They are most likely Muslim and never moved from the settlement they are currently living in. 30% of households comprise adults named with only Serbian names.[15] They expect that one extra year of schooling increases monthly income by roughly 17 euros, corresponding to 5% of the minimum wage.[16] Roma people usually do not

[11] A neighborhood is defined in the survey as an area corresponding to 200 square meters around their house. Therefore, settlements and neighborhoods may not exactly coincide. This is why we use both perceived ethnic composition of the neighbourhood and median perception in the settlement. In almost two-third of the settlements all households have the same perception. In the remaining third, there are only slight changes between *only Roma* and *mostly Roma*. We think that the median perception, by smoothing out potential outliers in household perceptions, may give a more accurate representation of the reality.

[12] In Serbia, primary school lasts eight years. There is a first cycle of four years, followed by a second cycle of other four years (Serbian Ministry of Education, Science and Technological Development, 2009).

[13] Data on monthly household income are inaccurate. A wealth index including durables and utilities and defined through the first principal component analysis is a better measure of the actual household wealth. Filmer and Pritchett (2001) showed that an index obtained through the first principal component can provide reasonable estimates of the wealth level effects in situations where wealth data are not directly available.

[14] Source of income is equal to 1 when the main source of income is a job in the formal sector, 2 when it is social benefits and 3 when it is a job in the informal sector.

[15] Examples of Serbian names are Aleksandar, Borislav, Ivan, Jelena, Katarina, Slobodan. Examples of Romani names are Alvin, Djemila, Djulijana, Ersijana, Nuredin, Roberto, Valentino. Romani names are different enough from Serbian names to clearly identify the ethnicity one belongs to (Behind the name, 2017).

[16] Expected returns to education were computed using questions in our survey about the salary parents expected for their children in different scenarios: no schooling, primary, secondary. We construct Mincerian expected returns by regressing log-income on years of education. We find that one more year of schooling increases expected log earning by 9.4% for boys and 8.1% for girls, in line with what is observed in the literature (Baudin et al., 2015; Jensen, 2010; Nguyen, 2008; Duflo, 2001; Montenegro and Patrinos, 2013; Hanushek and Welch, 2006).

perform jobs for which high levels of education are required. They mainly work in the informal sector, without written contracts, often self-employed especially in flea markets and more rarely in factories (LSMS 2007).[17]

[insert TABLE 1 here]

Columns (2), (3) and (4) report separately means for *only*, *mostly* and *few Roma* settlements. Households are overall comparable in terms of observable characteristics across settlement types. The differences in means are not statistically significant in almost all cases between *only* and *mostly*. More substantial differences are found between *only* or *mostly* and *few*. Wealth and share of women born in the settlement of residence are higher in *few Roma*, while number of adults as well as share of muslims are lower.

Panel B of Table 1 reports the characteristics of the settlements. Unfortunately the database of Roma settlements does not provide detailed information on all the settlements: our final sample counts 11 out of 13 settlements. Households' characteristics of these two not included settlements are not different from those of the other settlements. Households are equally located in urban and suburban areas, but *only Roma* settlements are more likely to be located in suburban areas.[18] The period when the first people settled in an area is classified in three categories: most of the settlements were originated between 1951 and 1972, but *only Roma* settlements are more likely to have been created more recently.[19] 33% of settlements were illegally created (1), 50% originated spontaneously in areas not regulated by the general urban plan (2), 8.3% were also spontaneously created but in areas where it was temporary permitted to build (3) and 8.3% originated spontaneously in allowed areas (4). *Only* and *mostly Roma* settlements are almost exclusively of the first two types, while *few Roma* settlements are evenly distributed across categories 2 and 3.

On average, Roma women in our sample have 3.2 children, of which 54% are boys. A preliminary investigation of our outcome of interest (Panel C of Table 1) shows that the number of children is significantly lower in *few Roma* settlements than in the other types of settlements. On average, in *few Roma* neighbourhoods there are 2.7 children per household, while in *mostly Roma* and in *only Roma* neighbourhoods there are respectively 3.2 and 3.6 children per house-

[17]More information on the Roma labour market in Serbia can be found in Battaglia and Lebedinski (2017).

[18]We define as urban area a local community with more than 35,000 inhabitants, in line with the definition of the Municipality of the City of Belgrade that distinguishes between urban and suburban areas on its own territory.

[19]The three categories defined by Society for Improvement of Roma settlements (2002) are the following: before 1944, 1951-1972, 1973-1995.

hold. The proportion of boys is not significantly different across the three groups, with slightly more boys in families in *few Roma* settlements.

3 Empirical Strategy and Results

3.1 OLS results

Summary statistics show that households in *only Roma* settlements have a higher fertility. Nonetheless they could be the reflection of different age structures, socio-economic conditions, family arrangements or returns to education. In this section, we test whether the gap in fertility across more or less segregated settlements persists once we take into account household and settlement characteristics. To do so, we estimate the following regression equation[20]:

$$F_{ijs} = \beta_0 + \beta_1 mostlyroma_s + \beta_2 fewroma_s + \gamma_1 X_{ijs} + \delta_1 S_s + \epsilon_{ijs} \quad (1)$$

where F_{ijs} stands for number of children for a woman i in household j in settlement s. $mostlyroma_s$ and $fewroma_s$ are dummies equals to 1 if the household j lives in a *mostly* or *few Roma* settlement, respectively, the omitted category being *only Roma*. An alternative measure of residential segregation at the street rather than at the settlement level is employed in Table D in the appendix.

The set of exogenous individual and household characteristics X_{ijs} and the vector of observable settlement characteristics S_s include controls for different mechanisms that could affect fertility. First, mother's age and age squared are present in all specifications, in order to make sure that our results are not driven by differences in the age structure of the female population across settlements. Columns (2) through (5) test each of the mechanisms mentioned in the introduction. In column (2), we include mothers' years of schooling, which is our closest proxy for the opportunity cost of female time, household wealth, as richer household may be able to afford having more children, and the main source of income (either from social benefits, informal or formal employment). In column (3) instead, we add a dummy for whether the mother was born in the settlement where she currently lives, in order to control for possible migration between settlements. In column (4), we use the number of adults in the household, to capture the fact that grandparents may help in taking care of larger cohorts of children, whether the

[20] We also use a Poisson model in order to take into account the fact that fertility is a count variable. Results are consistent and reported in Table D in the appendix.

settlement is in an urban or suburban area as a proxy for the cost of space and expected returns to education, which may influence the way parents allocate resources to quality versus quantity of children.[21] In column (5), we include cultural variables such as religion and whether parents' name are of Serbian origin.[22] The idea is that families that declare being of Christian Orthodox religion or whose first names are Serbian sounding might have been more influenced by Serbian cultural and social norms, among which that of having a small number of children.[23] Finally in column (6), we keep all controls that appeared to significantly affect fertility when taken separately and test their robustness.

Robust standard errors are clustered at the street level - the street where the household lives in.[24] Results are presented in Table 2.

[insert TABLE 2 here]

The first striking result is that fertility differences documented in the descriptive statistics persist once controlled for individual, household and settlement characteristics. In particular, women living in *mostly Roma* settlements seem to have about between a third and a quarter fewer children than those in *only Roma* locations, but the coefficient is not significant at the 10% level, and more importantly women in *few Roma* settlements have around 0.7 to 0.9 fewer children than similar women in *only Roma* locations, and that difference is significant at the 1% level in all specifications.

The second observation is that adding controls in columns (2) to (5) tends to reduce the fertility gap, while including all controls that came up significant reduces it by even more, as shown in column (6). Among significant controls, mothers' years of schooling and household wealth are lower the higher is fertility, suggesting that poorest households and households where mothers have a lower opportunity cost of time tend to have larger families. The source of income instead does not seem to matter. Mothers who were born in the same settlement tend to have significantly fewer children, which suggests that women who have migrated from their place of birth have larger families, possibly because they are the poorer on average.[25] Number of adults

[21]See the contributions by Kaufmann and Attanasio (2014); Jensen (2010); Nguyen (2008) for the impact of perceived returns to education on investment in education.

[22]More precisely, the religion dummy takes value 1 when Muslim and 0 when Christian Orthodox and other religions, but only 0.73% of our sample declares to practice another religion. We make use of the sounding of first names to capture acculturation in the spirit of recent papers like Algan et al. (2013); Abramitzky et al. (2016); Jurajda and Kovač (2016); Fouka (2017).

[23]In Serbia, 84% of the population is Christian Orthodox, 5% is Catholic, 3% is Muslim. The remaining 8% includes other religions, Atheists and people who do not declare their faith (Census, 2011).

[24]The results do not change when robust standard errors are clustered at the settlement level.

[25]Mothers who were born in their settlement of residence have a statistically significant higher wealth index

in the household and urban status barely affect fertility, while higher returns to education are strongly associated with a lower fertility, which illustrates the presence of a quality/quantity trade-off. Finally, religious affiliation does not seem to matter much, whereas exposure to the Serbian culture, as measured by whether parents hold Serbian sounding first names, is a strong predictor of a lower fertility.

When all significant controls are included together, all coefficients decrease in magnitude and some loose significance, illustrating that some of them are confounded. In particular, migration and mother's education are not significant anymore, suggesting in the first case that poorer household may self-select into migration producing a spurious relationship between migration and fertility, and in the second case that the education gradient in these Roma communities where very few women work in the formal sector does not matter drastically.

We investigate the robustness of these associations using alternative measures of segregation in Table E in the appendix. Results remain qualitatively consistent using the median perception about whether the neighbourhood is *only*, *mostly* or *few Roma* at the street level rather than at the settlement level. The same holds when we use the share of people who declare the neighbourhood is *only Roma* versus the rest.[26]

3.2 Birth timing

In the previous subsection, we have examined the number of children women have controlling for a second order polynomial in their age. It implicitly imposes a structured, though quite flexible, relationship between age and the number of children that is common to all women in the sample. Our conclusions are therefore valid for completed fertility if birth timing does not differ significantly across settlement types. In this subsection instead, we do away with this assumption by analyzing the pace at which women give birth instead of their total number of children only. For this, we use a proportional hazard model, which leaves the baseline hazard rate unspecified and assumes that it is shifted multiplicatively by covariates. We look at second and subsequent births and define the at-risk period as being 9 months after the previous birth we observe.[27] The at-risk period ends either with a birth or with the woman actually leaving

(p-value 0.0004). 71% of these women migrated from another neighborhood of Belgrade; 13% from another Serbian town and an additional 13% from Kosovo. The main reason to move is marriage. If we exclude from the analysis women who migrated, we obtain very similar results, which are not reported but available upon request.

[26] As reported in Table C in the appendix, the threshold parameter between *mostly Roma* and *few Roma* (*cut2*) appears to be not statistically different from zero, so the two categories can be collapsed together.

[27] The rationale for excluding first births is that it is complicated to define when the at-risk period starts, specially because we observe only children still in the household at the time of the survey, which may differ from

the sample (basically when time reaches her age at the time of the survey). Notice that now the unit of observation is the post birth spell and not the mother, hence the increase in the number of observations. Covariates are assumed to affect the baseline hazard multiplicatively, so that the equation we estimate is the following:

$$h(t|x_c) \quad = \quad h_0(t)e^{x_c\beta_x} \tag{2}$$

Where x_c includes the same controls as in the previous subsection, to which we add mother fixed effects in order to capture systematic variations in biological fecundity for instance. Robust standard errors are clustered at the street level.

[insert TABLE 3 here]

The baseline results shown in Table 3 confirm the findings from the previous subsection: women living in *few Roma* settlements have children at a significantly slower pace than their *mostly* or *only Roma* counterparts. The coefficient -0.569 on *few Roma* in column (1) corresponds to women in these neighbourhood being 43% less likely to have a child than comparable women living in an *only Roma* settlement. Controlling for potential mechanisms reduces but does not close the gap.[28]

Then we turn to the study of whether birth timing differs depending on the gender of the previously born. To this end, we interact our measure of segregation alternatively with a dummy indicating whether the previously born was a male or a female and another one indicating whether the first born was a female.[29] To make sure that children leaving the household early on does not introduce too much error in the measurement of the gender of the previously born or first born child, we choose to restrict the sample to mothers below 33 years of age. Results are presented in Table 4.[30]

the universe of all births. Indeed some of the oldest mothers in the sample may have had an early child who already left the sample and whom we consequently do not observe. Imbalances in the sex-ratio of first born children confirm our hypothesis, as reported in Figure B in the appendix.

[28]We test the proportional hazard assumption. As reported in Figure A in the appendix, the plotted curves are roughly parallel, providing evidence in its favor.

[29]This is related in spirit to a recent contribution by Dimri et al. (2017) who analyze the patterns of birth spacing in function of the gender of the next born, in the presence of sex-selective abortions.

[30]Results for the full sample are shown in Tables F and G in the appendix. Results remain qualitatively similar though statistical significance is lost, possibly due to the introduction of measurement error in the gender of previously born children. As reported in Figure C in the appendix, when we take only women below 33, we do not observe anymore an imbalanced sex-ratio of first born children, as is the case instead in Figure B for the full sample.

[insert TABLE 4 here]

The main result to highlight in this table is that the difference in birth spacing across settlement types comes mostly from the difference in birth spacing after a boy was born. Indeed, households in *only Roma* and *mostly Roma* do not exhibit significantly different spacing pattern after boys or girls. However, households living in a *few Roma* settlement tend to space substantially more after a boy, which is illustrated by the coefficient on *few Roma**male. The coefficients in column (1) indicate that the hazard ratio of having an extra child after a male is 23% smaller than after a girl for households in *only Roma* settlements (but this difference is not statistically different from zero), while it is 58% smaller in the case of *few Roma* households. This latter difference is significantly different from zero at the 5% level.

The same pattern emerges as controls are added: they allow to close only partially the gap in birth timing across settlement types. Results are qualitatively similar when using whether the first born is a male or a female. As before, we investigate the robustness of the analysis by using alternative measures of segregation. Results remain consistent, as reported in Table H in the appendix.

3.3 The endogeneity problem

The parameters of major interest in equation (1) are β_1 and β_2. Nonetheless, their identification is complicated as they may capture individual self-selection in particular areas rather than the effect of residential segregation itself. Living in one settlement or another is not random: fertility decisions and gender preferences can be related to the decision of living in a more or less segregated settlement. More educated and wealthier people tend to be more sensitive to family planning and have fewer children. If settlements that are more connected to the Serbian majority tend to attract educated individuals who care about their children's job opportunities and chase away the least educated, then some settlements may become less segregated and with a lower fertility because of the self-selection process but not because of segregation per se. Vice versa, if the wealthier in the Roma community tend to concentrate, leaving the poorest (and least able to migrate) more isolated from their community, then less segregated settlements would tend to be composed of poorer households who typically have a higher fertility. There is a potential selection bias into settlements whose direction is a priori unclear. More formally, in equation (1) the error term ϵ_{ijs} is composed of two parts:

$$\epsilon_{ijs} = \eta_{ijs} + u_{ijs}$$

where η_{ijs} is an unobservable individual term correlated with residential segregation and u_{ijs} is an unobservable individual term that is not correlated. $mostlyroma_s$ and $fewroma_s$ in (1) are likely to depend on some factors captured by η_{ijs}. A simple comparison of households in *only* or *mostly Roma* settlements versus *few* would therefore incorrectly estimate the effects of living in one or another area. In order to separate the effect of residential segregation from the impacts of the selection mechanism, we use an instrumental variable strategy. We need an instrument that is correlated with the current residential segregation but uncorrelated with the individual unobservable attributes affecting the fertility behavior.

The instrument for current residential segregation is (il)legal possibility to build in the area at the time of creation of the settlement. The dataset of Roma settlement (Society for Improvement of Roma settlements, 2002) provides information on when each settlement was created, whether it was forbidden or allowed to build at that time and how it formed, either spontaneously or illegally. The variable *(il)legal possibility to build* is categorical and corresponds to 1 if the settlement was created in areas where, according to the general urban plan, it was illegal to build, 2 if it was created spontaneously but in areas where there was no regulation allowing to build, 3 if it was created spontaneously with a temporary permit, and 4 if it was created spontaneously in allowed areas. The rationale behind is that, conditional on the year when it was created and its location (urban-suburban), a settlement where it was illegal to build at the time of its creation was more likely to become *only Roma* than a settlement created spontaneously with or without the municipality's permission. This is in line with the housing discrimination literature, which documents that it is more difficult for minorities to find accommodation both on the rental and the property markets. We are confident that the conditions of creation of settlements - which took place before any of the women in our sample were in childbearing age and which, for most of our sample, has rather involved their grand or great grandparents - do not affect current fertility, except through their influence on current residential segregation.

$typeofsettlement_s$ in a reduced form framework is therefore modelled as follows:

$$\begin{aligned} typeofsettlement_s &= \alpha_0 + \alpha_1 noregulation_s + \alpha_3 temporary_s + \\ &\quad + \alpha_3 allowed_s + \gamma_1 X_{ijs} + \delta_1 S_s + v_j \end{aligned} \tag{3}$$

where $noregulation_s$, $temporary_s$ and $allowed_s$ are dummy variables describing the whole universe of (il)legal possibility to build at the time of creation, the excluded category being "created in areas where it was illegal to build". Conditional on the year when the settlement was created and its location, the instrument affects the likelihood that settlements are currently segregated through past segregation but is unlikely to be correlated with any unobservable household or individual attributes that affect current fertility. Simple correlations reveal that the instrument is not significantly correlated with the observable characteristics of Roma people in each type of settlement.[31]

Table 5 shows the first stage with median perception about ethnic composition of the neighbourhood at the settlement level in column (1), at the street level in column (2) and the share of people declaring their neighbourhood is *only Roma* in column (3). As dependent variables in columns (1) and (2) are categorical, we chose to estimate these equations using ordered Probit, while specification (3) uses straight OLS. Table 5 confirms that the (il)legal possibility to build at the time of the creation of settlements is a good predictor of the level of residential segregation today. In the Hansen J-test of overidentifying restrictions, the null hypothesis is not rejected. The instrument set is appropriate: it satisfies the orthogonality condition.

[insert TABLE 5 here]

Of course, there exists some concerns with the instrument. Indeed, pre-existing characteristics of the area where a settlement originated could potentially capture the type of Roma the neighborhood attracts and these characteristics can be then transmitted over time. For instance, settlements created in areas where it was illegal to build may have attracted a less educated fringe of the Roma population. Human capital being both an important determinant of fertility and very transmissible over generations, we may confuse the effect of segregation with that of persistently low human capital. To mitigate this concern, we show, in Table 1, that in our sample, Roma in different types of settlements are comparable in terms of observable characteristics that affect fertility, such as educational attainment.

[31] It is only slightly negatively correlated with born in the same settlement and household wealth.

Furthermore, there is a temporal dimension that is potentially undesirable: the possibility to build in an area might simply capture ancient versus recent waves of migration to certain neighbourhoods or different attitudes towards illegal construction in the urban plan. We check that both illegal and legal settlements have appeared at all periods to discard that possibility.

3.4 IV results

Results for Limited Information Maximum Likelihood method are shown in Table 6. The baseline specification in column (1) suggests that living in an *only Roma* settlement increases fertility by roughly 1.5. This is a larger gap than estimated using straight OLS without instrumenting. It points at the existence of an attenuation bias in the way households self-select into different types of settlements. It gives credit to the hypothesis that less fertile and possibly richer households tend to concentrate and remain isolated, leaving the poorer and more fertile ones in less segregated areas.

Once we add controls for our different mechanisms, we obtain that the gap is decreased by the opportunity cost channel, left unchanged by the migration channel, increased when controlling for returns to education and almost completely closed when using cultural proxies. This is in line with what we obtain with simple OLS. Our interpretation is that while other mechanisms linked to socio-economic characteristics are not to disregard completely, exposure to the Serbian majority low fertility cultural norm has a first order effect and is the main explanatory variable of the fertility gradient in residential segregation.

Results hold when using alternative measures of segregation as shown in Table I in the appendix.

[insert TABLE 6 here]

3.5 Birth spacing patterns among Serbs

To give further evidence that a cultural transmission process is at work, we study the number of children and the birth spacing patterns among Serbs. Using data from the MICS 2010, we compute the average number of children among Serbs and show in Table J in the appendix that they do exhibit a longer spacing after boys than after girls, just like Roma living in *few Roma* settlements. It seems therefore that on top of a preference for smaller families - they have on average 1.73 children per woman, the Serbian majority also shares the same preference for sons rather than daughters with the Roma population most exposed to them.

While we are not aware of other studies documenting higher birth spacing after boys in Serbia or in other countries of Western Balkans, high sex ratios at birth have been recently observed in several countries of the region (UNFPA, 2012).[32] According to Guilmoto and Duthé (2013), together with the fertility decline and the development of modern healthcare services, the persistence of traditional patriarchal values is central to the son preference observed in these countries. Conversely, traditional Roma societies, especially in rural areas, show a female-biased sex ratio at birth and invest more heavily in daughters since they are more likely than sons to help their parents in taking care of siblings (Bereczkei and Dunbar, 1997, 2002).

4 Conclusion

We provide evidence that differences in fertility across neighborhoods within a given minority group depend on the concentration of the group one belongs to in that neighborhood. These differences remain even though socio-economic characteristics of their inhabitants are similar. Our data on the Roma populations from Belgrade in Serbia show that fertility in less segregated areas is lower than in more segregated areas and that women in *few Roma* settlements exhibit a higher preference for having at least one boy. While birth spacing is not significantly different across neighborhoods after girls, after boys women in less segregated areas tend to take much more time before the next birth. Our analysis of the potential mechanism points to the greater exposure of less segregated areas to the Serbian culture, in which fertility tends to be low and boys preferred.

In terms of policy recommendations, our results suggest that cultural inertia is an important factor to take into account while designing policies. Indeed, policies aiming at changing economic incentives of marginalized populations may remain poorly efficient for a while if these populations are severely segregated away from other fringes of the population. Policies promoting social mixing on the other hand could go a long way as exposure to different cultural norms seems conducive of rapid behavioral change.

[32] In Albania, the sex-ratio hovers around 110 (a normal sex ratio is 105). In Montenegro, for the period 2009-2011, the sex ratio at birth was 109.8. In Kosovo, for the period 2011-2013, the sex ratio at birth was 110.4. In Serbia, for the period 2000-2005, the sex ratio at birth was 107.

References

Abramitzky, R., L. P. Boustan, and K. Eriksson (2016). Cultural Assimilation during the Age of Mass Migration. Technical report, National Bureau of Economic Research.

Alesina, A., R. Baqir, and W. Easterly (1999). Public Goods and Ethnic Divisions. *The Quarterly Journal of Economics 114 (4)*, 1243–1284.

Algan, Y., T. Mayer, and M. Thoenig (2013). The Economic Incentives of Cultural Transmission: Spatial Evidence from Naming Patterns across France.

Battaglia, M. and L. Lebedinski (2015). Equal Access to Education: An Evaluation of the Roma Teaching Assistant Program in Serbia. *World Development 76*, 62–81.

Battaglia, M. and L. Lebedinski (2017). The Curse of Low Expectations:Remedial Education and Perceived Returns to Education of Roma People. *forthcoming in Economics of Transition*.

Baudin, T., D. de la Croix, and P. E. Gobbi (2015). Fertility and childlessness in the united states. *The American Economic Review 105*(6), 1852–1882.

Bayer, P., S. L. Ross, and G. Topa (2008). Place of Work and Place of Residence: Informal Hiring Networks and Labor Market Outcomes. *Journal of Political Economy 116 (6)*, 1150–1196.

Behind the name (2017). `http://www.behindthename.com/names/usage`.

Bereczkei, T. and R. I. M. Dunbar (1997). Female-biased Reproductive Strategies in a Ethnic Hungarian Gypsy Population. *Royal Society Proceedings, Biological Sciences 264*, 17–22.

Bereczkei, T. and R. I. M. Dunbar (2002). Helping-at-the-Nest and Sex-Biased Parental Investment in a Hungarian Gypsy Population. *Current Anthropology 43*(5), 804–809.

Blau, F., L. Kahn, A. Liu, and K. Papps (2013). The transmission of women?s fertility, human capital, and work orientation across immigrant generations. *Journal of Population Economics 26*(2), 405–435.

Blume, L. and S. Durlauf (2006). *Identifying Social Interactions: A Review.* Michael Oakes and Jay Kaufman (eds.), Methods in Social Epidomology, Wiley.

Boeri, T., M. D. Philippis, E. Patacchini, and M. Pellizzari (2015). Immigration, Housing Discrimination and Employment. *The Economic Journal 125 (508)*, F82–F114.

Boustan, L. P. (2012). School Desegregation and Urban Change: Evidence from City Boundaries. *American Economic Journal: Applied Economics 4*(1), 85–108.

Boustan, L. P. and R. A. Margo (2013). A Silver Lining to White Flight? White Suburbanization and African–American Homeownership, 1940–1980. *Journal of Urban Economics 78*, 71–80.

Card, D. and J. Rothstein (2007). Racial Segregation and the Black-White Test Score Gap. *Journal of Public Economics 91 (11-12)*, 2158–2184.

Cervellati, M. and U. Sunde (2015). The Economic and Demographic Transition, Mortality, and Comparative Development. *American Economic Journal: Macroeconomics 7*(3), 189–225.

Chabé-Ferret, B. (2013). Socioeconomic Characteristics, Fertility Norms and the Black-White Fertility Gap in the US. Technical report, Université catholique de Louvain, Institut de Recherches Economiques et Sociales (IRES).

Chabé-Ferret, B. (2016). Adherence to Cultural Norms and Economic Incentives: Evidence from Fertility Timing Decisions.

Chabé-Ferret, B. and P. Melindi Ghidi (2013, July). Differences in Fertility Behavior and Uncertainty: an Economic Theory of the Minority Status Hypothesis. *Journal of Population Economics 26*(3), 887–905.

Clark, K. and S. Drinkwater (2002). Enclaves, Neighbourhood Effects and Employment Outcomes: Ethnic Minorities in England and Wales. *Journal of Population Economics 7*(5), 603–628.

Cutler, D. M. and E. L. Glaeser (1997). Are Ghettos Good or Bad? *The Quarterly Journal of Economics, 112 (3)*, 827–872.

de la Croix, D., F. Perrin, et al. (2016). French Fertility and Education Transition: Rational Choice vs. Cultural Diffusion. *IRES Discussion Papers* (2016-7), 36.

Dimri, A., V. Gille, and P. Ketz (2017). Sex-selective Abortion: Are There Repeated Abortions? *Mimeo*.

Doepke, M. (2015). Gary Becker on the Quantity and Quality of Children. *Journal of Demographic Economics 81*(1), 5966.

Duflo, E. (2001). Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment. *The American Economic Review 91*(4), 795–813.

Durlauf, S. (2004). Chapter 50 Neighborhood Effects. In J. V. Henderson and J.-F. Thisse (Eds.), *Cities and Geography*, Volume 4 of *Handbook of Regional and Urban Economics*, pp. 2173–2242. Elsevier.

Edin, P.-A., P. Frederiksson, and O. Åslund (2003). Ethnic Enclaves and the Economic Success of Immigrants: Evidence from a Natural Experiment. *The Quarterly Journal of Economics 118*(1), 329–357.

Fernández, F. and A. Fogli (2006). Fertility: The Role of Culture and Family Experience. *Journal of the European Economic Association 4 (2-3)*, 552–561.

Fernández, F. and A. Fogli (2009). Culture: An Empirical Investigation of Beliefs, Work and Fertility. *American Economic Journal: Macroeconomics 1 (1)*, 146–77.

Filmer, D. and L. Pritchett (2001). Estimating Wealth Effects Without Expenditure Data or Tears: An Application to Educational Enrollments in States of India. *Demography 38 (1)*, 115–132.

Fouka, V. (2017). How do Immigrants respond to Discrimination? The Case of Germans in the US during World War I.

Galor, O. (2012, January). The Demographic Transition: Causes and Consequences. *Cliometrica, Journal of Historical Economics and Econometric History 6*(1), 1–28.

Galor, O. and D. N. Weil (1999). From Malthusian Stagnation to Modern Growth. *The American Economic Review 89*(2), 150–154.

Galor, O. and D. N. Weil (2000). Population, Technology, and Growth: From Malthusian Stagnation to the Demographic Transition and Beyond. *American economic review*, 806–828.

Gobbi, P. E. (2013). A Model of Voluntary Childlessness. *Journal of Population Economics 26*(3), 963–982.

Goldscheider, C. and P. R. Uhlenberg (1969). Minority Group Status and Fertility. *American Journal of Sociology 74*(4), 361–372.

Guilmoto, C. Z. and G. Duthé (2013). Masculinization of Birth in Eastern Europe. *Population and Societies* (506).

Hanushek, E. A. and F. Welch (2006). *Handbook of the Economics of Education*. North-Holland.

Jaksic, B. and G. Basic (2010). Roma Settlements, Living Conditions and the Possibility for Integration of Roma in Serbia. Ethnicity Research Centre.

Jensen, R. (2010). The (perceived) Returns to Education and the Demand for Schooling. *The Quarterly Journal of Economics 125 (2)*, 515–548.

Jurajda, S. and D. Kovač (2016). What's in a name in a war.

Kalemli-Ozcan, S. (2002). Does the Mortality Decline promote Economic Growth? *Journal of Economic Growth 7*(4), 411–439.

Kaufmann, K. M. and O. Attanasio (2014). Educational Choices and Returns to Schooling: Mothers' and Youths' Subjective Expectations and Their Role by Gender. *Journal of Development Economics 109(C)*, 201–216.

Klemp, M. and J. L. Weisdorf (2012). Fecundity, Fertility and Family Reconstitution Data: The Child Quantity-Quality Trade-Off Revisited. *CEPR Discussion Papers 9121*.

La Ferrara, E. and A. Mele (2006). Racial Segregation and Public School Expenditure. *CEPR Discussion Paper 5750*.

Li, H. and J. Zhang (2007). Do High Birth Rates hamper Economic Growth? *The Review of Economics and Statistics 89*(1), 110–117.

Manley, D., M. van Ham, and J. Doherty (2011). Social Mixing as a Cure for Negative Neighbourhood Effects: Evidence Based Policy or Urban Myth? In T. B. . L. L. Bridge, G. (Ed.), *Mixed Communities: Gentrification by Stealth*. Policy Press, Bristol.

Montenegro, C. E. and H. A. Patrinos (2013). Returns to Schooling around the World. Technical report, Background paper for the World Development Report 2013.

Nguyen, T. (2008). *Information, Role Models and Perceived Returns to Education: Experimental Evidence from Madagascar.* Ph. D. thesis, Massachusetts Institute of Technology.

Open Society Institute (2007). Equal Access to Quality Education For Roma. Technical report, Open Society Foundation.

Serbian Ministry of Education, Science and Technological Development (2009). Law on the foundations of the education system in serbia.

Society for Improvement of Roma settlements (2002). Overview of Roma settlements, Belgrade. Technical report, Society for Improvement of Roma settlements.

Stichnoth, H. and M. Yeter (2016, September). Cultural Influences on the Fertility Behavior of First- and Second-Generation Immigrants. *JODE - Journal of Demographic Economics 82*(3), 281–314.

UNFPA (2012). Sex Imbalances at Birth: Current Trends, Consequences and Policy Implications. Technical report, UNFPA Asia and the Pacific Regional Office.

5 Tables

Table 1: Households' and Settlements' Characteristics

Variable	All	Only Roma Settlements	Mostly Roma Settlements	Few Roma Settlements	(3)-(2)	(4)-(2)	(4)-(3)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
A. Households' characteristics							
mother's age	32.500	32.297	32.435	33.609	.018	.185	.146
	(5.666)	(4.799)	(6.092)	(5.220)			
mother's years of schooling	5.360	5.459	5.333	5.391	-.033	-.014	.018
	(3.141)	(3.449)	(2.948)	(3.539)			
household wealth[a]	.014	-.097	-.015	.567	.036	**.281**	**.265**
	(1.567)	(1.720)	(1.476)	(1.627)			
household source income (=1 if formal sector job)	.492	.527	.478	.478	-.069	-.068	.000
	(.501)	(.503)	(.501)	(.511)			
household source income (=1 if social benefits)	.167	.149	.161	.261	.025	.195	.171
	(.373)	(.358)	(.369)	(.449)			
household source income (=1 if informal sector job)	.341	.324	.360	.261	.053	-.097	-.151
	(.475)	(.471)	(.482)	(.449)			
number of adults	2.244	2.500	2.168	1.957	**-.268**	**-.433**	-.215
	(.817)	(1.037)	(.682)	(.706)			
muslim (=1)	.736	.865	.708	.522	**-.274**	**-.557**	**-.272**
	(.441)	(.344)	(.456)	(.511)			
mother born in the same settlement (=1)	.767	.716	.770	.913	.087	**.366**	**.280**
	(.423)	(.454)	(.422)	(.288)			
only serbian names (adults) (=1)	.289	.203	.303	.478	.098	.132	.035
	(.280)	(.268)	(.267)	(.312)			
expected returns to education in dinars (street)[b]	1709	1681	1720	1727	.040	.050	.010
	(621)	(770)	(561)	(496)			
Obs.	258	74	161	23			
B. Settlements' characteristics							
urban (=1)	.417	.333	.429	.500	.121	.183	.081
	(.515)	(.577)	(.535)	(.707)			
illegal	.333	.333	.429	0	.121	**-.577**	**-.802**
	(.492)	(.577)	(.535)	(.)			
spontaneous, unregulated	.5	.667	.429	.5	**-.303**	-.183	.081
	(.522)	(.577)	(.535)	(.707)			
spontaneous, temporary permit	.083	0	0	.5	0	**.707**	**.707**
	(.289)	(.)	(.)	(.707)			
spontaneous, allowed[c]	.083	0	.143	0	**.378**	0	**-.378**
	(.289)	(.)	(.378)	(.707)			
year of creation (before 1944)	.25	.333	.144	.5	**-.276**	.183	**.445**
	(.452)	(.577)	(.378)	(.707)			
year of creation (1951-1972)	.333	0	.428	.5	**.802**	**.707**	.081
	(.492)	(.)	(.534)	(.707)			
year of creation (1973-1995)	.417	.667	.428	0	**-.303**	**-1.155**	**.802**
	(.515)	(.577)	(.534)	(.)			
Obs.	12	3	7	2			
C. Fertility outcomes							
number of children	3.263	3.581	3.199	2.696	-0.207	**-0.511**	**-0.332**
	(1.250)	(1.434)	(1.161)	(0.974)			
proportion of boys	0.538	0.495	0.551	0.578	0.148	0.192	0.062
	(0.275)	(0.269)	(0.267)	(0.340)			
Obs.	258	74	161	23			

Columns (5), (6) and (7) represent normalized differences.
[a] The wealth index ranges between -3.135 and 2.865.
[b] 1709 dinars correspond to roughly 17 euro (1 RSD = 0.009626 Euro, November 2011).
[c] Pearson's χ^2 test of dissimilarity of distributions is equal to 841.8353. Pr = 0.000

Table 2: Number of children - OLS

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS
mostly roma	-0.345	-0.322**	-0.319	-0.311	-0.267	-0.262*
	(0.225)	(0.158)	(0.212)	(0.205)	(0.188)	(0.148)
few roma	-0.912***	-0.785***	-0.805***	-0.868***	-0.811***	-0.709***
	(0.277)	(0.201)	(0.275)	(0.267)	(0.273)	(0.185)
mother's years of schooling		-0.069**				-0.048
		(0.030)				(0.031)
household wealth		-0.202***				-0.166***
		(0.052)				(0.047)
source income (social benefit)		0.154				
		(0.211)				
source income (informal sector job)		0.209				
		(0.191)				
same settlement			-0.479**			-0.213
			(0.209)			(0.226)
number of adults				0.047		
				(0.104)		
urban				-0.013		
				(0.173)		
expected return to education				-0.346***		-0.239**
				(0.112)		(0.098)
muslim					0.152	
					(0.272)	
only serbian names					-0.883**	-0.595**
					(0.335)	(0.230)
Mother's age and age squared	x	x	x	x	x	x
Obs.	258	258	258	258	258	258
r2	0.056	0.191	0.080	0.085	0.132	0.224

Robust standard errors clustered at the street level in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3: Birth spacing

	(1) COX	(2) COX	(3) COX	(4) COX	(5) COX	(6) COX
mostly roma	-0.157	-0.152	-0.159	-0.126	-0.114	-0.150
	(0.126)	(0.096)	(0.128)	(0.131)	(0.123)	(0.104)
few roma	-0.569***	-0.534***	-0.581***	-0.518***	-0.503***	-0.549***
	(0.131)	(0.100)	(0.130)	(0.155)	(0.118)	(0.095)
Mother's age and age squared	x	x	x	x	x	x
Opportunity cost of time ch.		x				x
Migration ch.			x			x
Socio-demographic ch.				x		x
Cultural ch.					x	x
Obs.	836	836	836	836	836	836

Robust standard errors clustered at the street level in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4: Birth Spacing if woman aged less than 33

	(1) COX	(2) COX	(3) COX	(4) COX	(5) COX	(6) COX	(7) COX	(8) COX	(9) COX	(10) COX	(11) COX	(12) COX
mostly roma	-0.092	-0.040	-0.081	-0.111	-0.092	-0.040	-0.061	-0.008	-0.069	0.009	-0.085	-0.098
	(0.143)	(0.188)	(0.125)	(0.176)	(0.142)	(0.188)	(0.157)	(0.213)	(0.143)	(0.159)	(0.132)	(0.181)
few roma	-0.167	-0.704***	-0.012	-0.477**	-0.165	-0.703***	-0.103	-0.653***	-0.166	-0.495***	-0.161	-0.490***
	(0.152)	(0.213)	(0.173)	(0.223)	(0.154)	(0.210)	(0.138)	(0.246)	(0.187)	(0.190)	(0.191)	(0.162)
male	-0.259		-0.234		-0.262		-0.264		-0.244		-0.183	
	(0.180)		(0.185)		(0.176)		(0.183)		(0.181)		(0.178)	
mostly roma*male	0.098		0.023		0.099		0.096		0.100		0.026	
	(0.238)		(0.254)		(0.237)		(0.238)		(0.238)		(0.250)	
few roma*male	-0.607**		-0.585**		-0.603**		-0.633**		-0.505**		-0.513**	
	(0.274)		(0.289)		(0.271)		(0.279)		(0.225)		(0.226)	
firstbirth girl		0.197		0.162		0.197		0.189		0.197		0.142
		(0.172)		(0.190)		(0.170)		(0.178)		(0.175)		(0.189)
mostly roma*firstbirth girl		-0.069		0.027		-0.069		-0.088		-0.125		0.011
		(0.230)		(0.251)		(0.230)		(0.226)		(0.220)		(0.251)
few roma*firstbirth girl		0.163		0.070		0.163		0.140		-0.067		-0.035
		(0.264)		(0.304)		(0.262)		(0.291)		(0.197)		(0.191)
Obs.	531	531	531	531	531	531	531	531	531	531	531	531

Robust standard errors clustered at the street level in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. In columns (1) to (2) we only control for mother's age and age squared. In columns (3) to (4) we also control for opportunity cost of time characteristics; in columns (5) and (6) for whether the mother was born in the same settlement; in columns (7) and (8) for socio-demographic characteristics and in columns (9) and (10) for cultural characteristics. In columns (11) and (12) we control for mother's age and age squared, mother's years of schooling, household wealth, expected returns to education and whether the parents have got only serbian names.

Table 5: First stage - *Type of settlement*

	Type Settlements OPROBIT (1)	Type Street OPROBIT (2)	Share Only Roma Settlement OLS (3)
spontaneous w/out regulation	-0.196	0.266	-24.474***
	(0.565)	(0.480)	(8.804)
spontaneous with temporary permit	7.091***	2.028***	-107.838***
	(0.490)	(0.783)	(8.827)
spontaneous with permission	0.576	0.374	-104.183***
	(0.389)	(0.342)	(8.894)
Obs.	258	258	258
Test of overidentifying restriction (p-value)	0.4964	0.8712	0.3529

Robust standard errors clustered at the street level in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Column (1) has as a dependent variable the categorical indicator capturing exactly whether the settlement is of *few Roma, mostly Roma or only Roma.* Column (2) has as a dependent variable the categorical indicator capturing exactly whether the street is of *few Roma, mostly Roma or only Roma.* Column (3) has as a dependent variable the endogenous continuous variable *share of only roma.*

Table 6: Number of children - IV

	(1) IV	(2) IV	(3) IV	(4) IV	(5) IV	(6) IV
mostly roma	-0.984	-0.582	-1.018	-1.301	0.381	-0.070
	(0.857)	(0.643)	(0.809)	(1.116)	(0.900)	(0.679)
few roma	-1.447*	-0.964*	-1.430**	-1.862**	-0.121	-0.564
	(0.741)	(0.575)	(0.721)	(0.934)	(0.689)	(0.603)
mother's years of schooling		-0.070**				-0.046
		(0.028)				(0.033)
household wealth		-0.196***				-0.167***
		(0.058)				(0.046)
source income (social benefit)		0.170				
		(0.210)				
source income (informal sector job)		0.223				
		(0.185)				
same settlement			-0.421*			-0.228
			(0.224)			(0.247)
number of adults				-0.070		
				(0.157)		
urban				0.125		
				(0.246)		
expected return to education				-0.303*		-0.243***
				(0.175)		(0.093)
muslim					0.307	
					(0.232)	
only serbian names					-0.840***	-0.614**
					(0.322)	(0.241)
Mother's age and age squared	x	x	x	x	x	x
Obs.	258	258	258	258	258	258

Robust standard errors clustered at the street level in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

6 Figures

Figure 1: Map of Belgrade with settlements.

A Appendix

Table A: Characteristics of Roma versus Non-Roma

Demographic characteristics	Roma	Non-Roma
Household size	5.60	3.50
Number of children younger than 18 (in household)	2.40	0.86
Age (average)	24.75	35.65
Population younger than 18	0.43	0.25
Employment rate* (males)	0.57	0.56
Employment rate* (females)	0.11	0.40
Individuals below the poverty line*	0.46	0.07
Education		
Children between 6 and 15 not enrolled in school	0.11	0.01
Unfinished primary school	0.29	0.04

Source: Serbia - MICS 2010.

* Source: Serbia - LSMS 2007

Table B: Households' Characteristics - Comparison

	Our data	MICS 2010	
Variable	Roma	Roma	Non-Roma
	(1)	(2)	(3)
Households' characteristics			
age	32.506	28.12	31.84
	(5.66)	(5.71)	(4.56)
age at first birth	20.15	20.12	26.10
	(4.07)	(3.82)	(4.42)
mother's years of schooling	5.36	7.46	12.87
	(3.14)	(2.47)	(2.16)
number of adults	2.24	3.45	2.98
	(.82)	(1.47)	(1.23)
born in Serbia* (=1)	.844	.895	.907
Obs.	274	262	528

* Source: Serbia - LSMS 2007

Table C: Threshold parameters for categories

	Coef.	Robust Std. Err.	95% Conf. Interval	
$/cut1$	-22.237	4.008	-30.093	-14.380
$/cut2$	-7.169	3.967	-14.944	0.605

Table D: Number of children - POISSON

	(1) POISSON	(2) POISSON	(3) POISSON	(4) POISSON	(5) POISSON	(6) POISSON
mostly roma	-0.102	-0.091**	-0.094	-0.091	-0.077	-0.069
	(0.064)	(0.044)	(0.061)	(0.058)	(0.053)	(0.044)
few roma	-0.292***	-0.251***	-0.260***	-0.278***	-0.260***	-0.224***
	(0.089)	(0.065)	(0.089)	(0.086)	(0.086)	(0.063)
mother's years of schooling		-0.021**				-0.014
		(0.008)				(0.009)
household wealth		-0.061***				-0.050***
		(0.015)				(0.013)
source income (social benefit)		0.048				
		(0.063)				
source income (informal sector job)		0.063				
		(0.058)				
same settlement			-0.138**			-0.056
			(0.058)			(0.065)
number of adults				0.014		
				(0.031)		
urban				-0.004		
				(0.051)		
expected return to education				-0.103***		-0.068**
				(0.032)		(0.029)
muslim					0.049	
					(0.085)	
only serbian names					-0.304***	-0.218**
					(0.115)	(0.086)
Mother's age and age squared	x	x	x	x	x	x
Obs.	258	258	258	258	258	258

Robust standard errors clustered at the street level in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table E: Number of children: Alternative measures of residential segregation

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS
mostly roma street	-0.592***	-0.469***	-0.557***	-0.473**	-0.462***	-0.349**
	(0.173)	(0.141)	(0.162)	(0.190)	(0.140)	(0.152)
few roma street	-1.006***	-0.700***	-0.918***	-0.864***	-0.851***	-0.568***
	(0.194)	(0.187)	(0.186)	(0.208)	(0.190)	(0.190)
share only roma	0.010***	0.008***	0.009***	0.008*	0.008**	0.006**
	(0.004)	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)
Mother's age and age squared	x	x	x	x	x	x
Opportunity cost of time ch.		x				x
Migration ch.			x			x
Socio-demographic ch.				x		x
Cultural ch.					x	x
Obs.	258	258	258	258	258	258

Robust standard errors clustered at the street level in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table F: Birth spacing - gender differences using previously born children

	(1) COX	(2) COX	(3) COX	(4) COX	(5) COX	(6) COX
mostly roma	-0.168	-0.172*	-0.165	-0.133	-0.149	-0.143
	(0.125)	(0.099)	(0.127)	(0.137)	(0.122)	(0.101)
few roma	-0.287**	-0.304***	-0.289**	-0.219**	-0.301**	-0.311***
	(0.128)	(0.101)	(0.132)	(0.103)	(0.143)	(0.115)
male	-0.281*	-0.297*	-0.273*	-0.271*	-0.304*	-0.258*
	(0.160)	(0.157)	(0.157)	(0.151)	(0.158)	(0.148)
mostly roma*male	0.061	0.078	0.051	0.051	0.099	0.030
	(0.191)	(0.190)	(0.191)	(0.187)	(0.190)	(0.187)
few roma*male	-0.415	-0.332	-0.424	-0.445	-0.294	-0.348
	(0.281)	(0.253)	(0.282)	(0.272)	(0.223)	(0.234)
Mother's age and age squared	x	x	x	x	x	x
Opportunity cost of time ch.		x				x
Migration ch.			x			x
Socio-demographic ch.				x		x
Cultural ch.					x	x
Obs.	836	836	836	836	836	836

Robust standard errors clustered at the street level in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table G: Birth spacing - gender differences using first born

	(1) COX	(2) COX	(3) COX	(4) COX	(5) COX	(6) COX
mostly roma	-0.208	-0.205	-0.208	-0.175	-0.151	-0.200
	(0.179)	(0.141)	(0.183)	(0.185)	(0.172)	(0.154)
few roma	-0.691***	-0.628***	-0.700***	-0.639**	-0.590***	-0.618***
	(0.223)	(0.160)	(0.222)	(0.256)	(0.162)	(0.146)
firstbirth girl	0.094	0.141	0.095	0.075	0.096	0.131
	(0.148)	(0.130)	(0.151)	(0.130)	(0.147)	(0.131)
mostly roma*firstbirth girl	0.089	0.092	0.085	0.075	0.052	0.079
	(0.206)	(0.195)	(0.211)	(0.193)	(0.203)	(0.202)
few roma*firstbirth girl	0.257	0.196	0.256	0.233	0.165	0.131
	(0.285)	(0.212)	(0.292)	(0.296)	(0.207)	(0.209)
Mother's age and age squared	x	x	x	x	x	x
Opportunity cost of time ch.		x				x
Migration ch.			x			x
Socio-demographic ch.				x		x
Cultural ch.					x	x
Obs.	836	836	836	836	836	836

Robust standard errors clustered at the street level in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table H: Birth Spacing - Alternative measures of residential segregation

	(1) COX	(2) COX	(3) COX	(4) COX	(5) COX	(6) COX
mostly roma street	-0.208	-0.205	-0.208	-0.175	-0.151	-0.200
	(0.179)	(0.141)	(0.183)	(0.185)	(0.172)	(0.154)
few roma street	-0.691***	-0.628***	-0.700***	-0.639**	-0.590***	-0.618***
	(0.223)	(0.160)	(0.222)	(0.256)	(0.162)	(0.146)
share onlyroma	0.005**	0.005***	0.005**	0.005*	0.004**	0.005**
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Obs.	836	836	836	836	836	836

Robust standard errors clustered at the street level in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. In columns (1) we only control for mother's age and age squared. In columns (2) we also control for opportunity cost of time characteristics; in columns (3) for whether the mother was born in the same settlement; in columns (4) for socio-demographic characteristics and in columns (5) for cultural characteristics. In columns (6) we control for mother's age and age squared, mother's years of schooling, household wealth, expected returns to education and whether the parents have got only serbian names.

Table I: Number of children: Alternative measures of residential segregation - IV

	(1) IV	(2) IV	(3) IV	(4) IV	(5) IV	(6) IV
mostly roma street	-1.944*	-1.344	-1.961*	-2.069	-0.266	-0.667
	(1.161)	(1.031)	(1.104)	(1.301)	(2.189)	(0.956)
few roma street	-2.427**	-1.731*	-2.412**	-2.488**	-0.857	-1.143
	(1.060)	(0.963)	(1.057)	(1.081)	(2.325)	(0.871)
share only roma	0.009*	0.004	0.006	0.005	0.005	0.001
	(0.005)	(0.004)	(0.005)	(0.005)	(0.005)	(0.005)
Obs.	258	258	258	258	258	258

Robust standard errors clustered at the street level in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. In columns (1) we only control for mother's age and age squared. In columns (2) we also control for opportunity cost of time characteristics; in columns (3) for whether the mother was born in the same settlement; in columns (4) for socio-demographic characteristics and in columns (5) for cultural characteristics. In columns (6) we control for mother's age and age squared, mother's years of schooling, household wealth, expected returns to education and whether the parents have got only serbian names.

Table J: Birth spacing - gender differences using previously born children - Serbs

	(1) COX	(2) COX	(3) COX
average number of children: 1.73 (0.737)			
male	-0.100***	-0.111**	-0.132***
	(0.042)	(0.043)	(0.047)
age mother		0.057	0.170**
		(0.061)	(0.071)
age mother sq		-0.002**	-0.004***
		(0.001)	(0.001)
Obs.	4701	4701	4701

Robust standard errors clustered at the district level in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. In column (1) no additional controls are added; in column (2) we only control for mother's age and age squared; in column (3) we also control for the ranking among children, mother's level of education, wealth and district fixed effects. Source: MICS 2010.

Figure A: Test of the proportional-hazards assumption

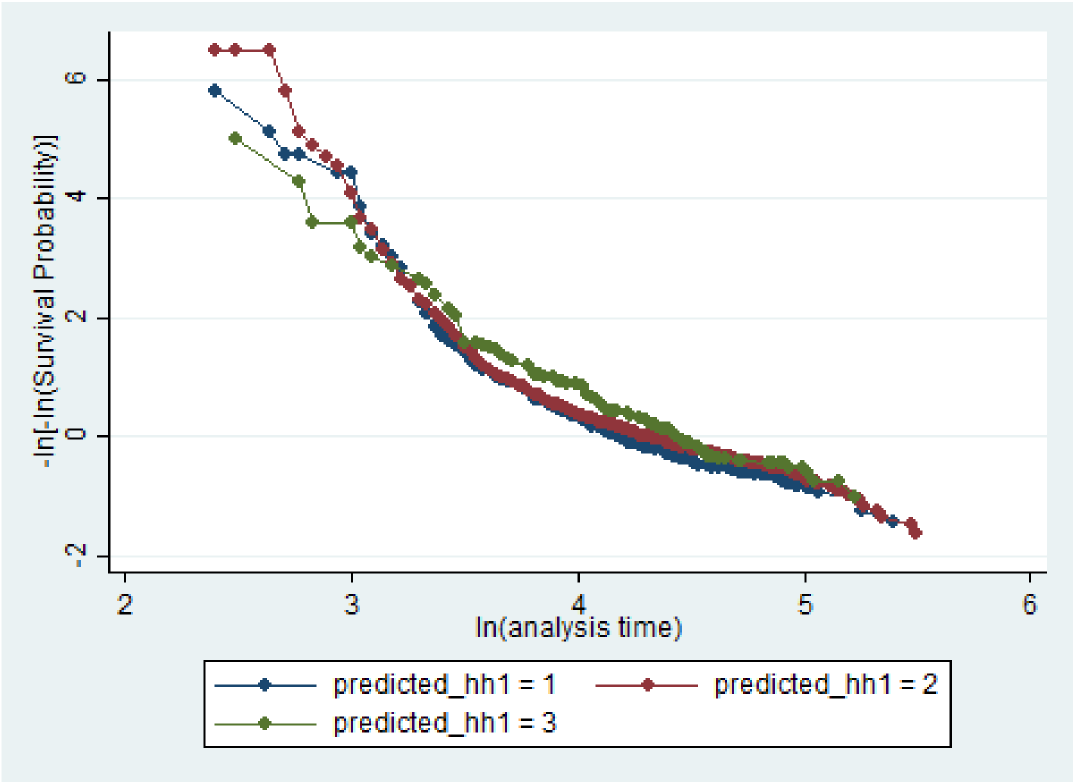

Figure B: Sex ratio by birth and settlement

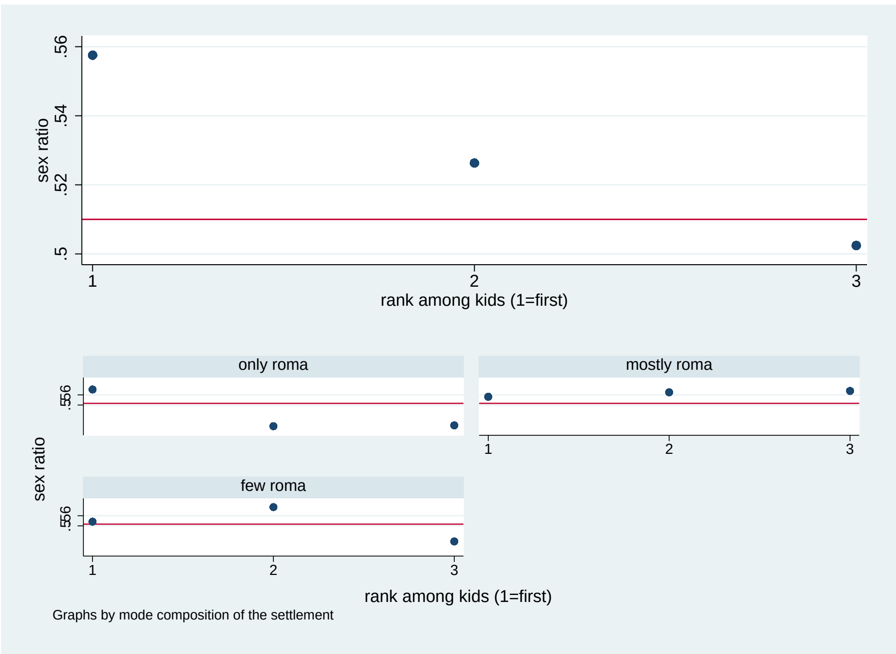

Figure C: Sex ratio by birth and settlement if women is aged less than 33