The Economic Incentives of Cultural Transmission: Spatial Evidence from Naming Patterns across France*

Yann Algan† Thierry Mayer‡ Mathias Thoenig§

January 25, 2013

Abstract

This paper estimates a model of cultural transmission by looking at the type of first names given by parents to their newborn children. We focus on the transmission of Arabic name versus Non-Arabic name in the French society and disentangle the relative contribution of vertical transmission of parental culture, horizontal transmission from the neighborhood's culture and the economic penalty associated with names that sound culturally distinctive in the French society. We use the French Labor Force Survey which is based on data collection of names and socio-economic characteristics for all individuals living within the same residential block of 20 neighboring households. To get rid of endogenous residential sorting, our estimates are based on the sample of households who are randomly allocated across public housings dwellings. We find that parents do take into account the expected economic cost that they inflict to their child by choosing a culturally distinctive name. Our estimates imply that the parents are ready to abandon one year of expected average income for their child in order to transmit their cultural trait.

*We thank Alberto Alesina, Alberto Bisin, Paola Giuliano, Laurent Gobillon, Yves Zenou, Fabrizio Zilibotti, for helpful comments. We also thank participants to seminars and conferences at RSAI in Miami, the NBER Summer Institute in Boston, CEPR conference on culture and economics in Rome, INED in Paris, UCD in Dublin, UAB and Pompeu Fabra in Barcelona, CIFAR IAST at the Toulouse School of Economics, Universities of Bern, Lausanne, and Lugano, for their helpful comments. Yann Algan acknowledges financial support from the ERC Starting Grant 240923 under the European Community Seventh Framework Programme (FP7/2007-2013). Mathias Thoenig acknowledges financial support from the ERC Starting Grant GRIEVANCES-313327.

†Sciences Po and CEPR.

‡Sciences Po, CEPII and CEPR.

§University of Lausanne and CEPR.

1 Introduction

In this paper we estimate a model of inter-generational cultural transmission at the household level. The cultural trait under scrutiny corresponds to the type of first name given by parents to their newborn children. We focus on the desire of parents to transmit their own trait, potentially conflicting with peers' influence, and the adverse impact that a very distinct name can have on children future wage/employment status. Our main contribution is to estimate the extent to which parents are willing to allow for a deterioration of expected economic outcomes (for the child) in order to satisfy their desire to transmit their own trait, or to conform to the dominant trait in the surrounding area. We consider this analysis as a first attempt to identify the economic determinants of cultural transmission, while the previous literature has focused so far on the effect of culture on economic outcomes (see Fernandez, 2008).

The analysis of naming decisions is an ideal experiment to probe into the interplay between economic incentives and cultural transmission. The first name is a crucial marker of cultural identity. As stressed by sociologists (Lieberson, 2000) the choice of first names are available to all parents, without material constraints, and thus are a pure expression of cultural identity. Levitt and Fryer (2004) have provided additional evidence on the cultural component of first name by showing that the surge in distinctively Black names in the US since the seventies could be associated to a rise in the Black cultural identity. But the first name is also perceived as a signal for the employer about the cultural and socio-economic background of the individual and can have direct economic consequence. In particular, various audit studies show that first names associated to a cultural minority are perceived negatively by employers (see the seminal study by Bertrand and Mullainathan, 2004). We add to this literature by estimating to what extent parents do take into account this economic cost in their naming decision.

We start by providing a theoretical model of naming decision. We incorporate the two traditional vertical and horizontal channels analyzed in the literature on cultural transmission (see Bisin and Verdier, 2001), but include on top of that an economic channel. The vertical transmission channel results from the utility gain for *parents* of transmitting their own cultural type; the horizontal transmission channel results from spatial externalities associated to the cultural type of peers and neighbors; and lastly the economic channel results from the utility gains/losses for *children* linked to the expected economic outcomes associated with their cultural type.

In a second stage we estimate our structural model of naming decision. Our dependent variable is the type/cultural category of the first names of babies born in France between 2003 and 2007. We specifically focus on the transmission of Arabic names as opposed to non-Arabic names in the French society. Non-Arabic Names in France are mainly associated with Saint Names that come from the French calendar of Saints. They are deeply associated with the French culture and are typically hold by natives whose both parents and grand-parents were born in France. By opposition the Arabic names are associated with the most important population of immigrants in France after World War Two, in particular in the aftermath of the decolonization initiated in the 1960s. Those names capture a cultural heritage that is potentially the most conflicting with that of the French

culture. First they are to a large extent a signal of the Muslim religious affiliation since most of those names come from the Qu'ran, and the transmission of first names associated with the Qu'ran is a natural practice for religious people. Second, they are associated with a hatred decolonization process in the French history, exemplified by the war of Algeria (1952-1964). XXX changer le phrasing pour limiter la repetition avec section XXX Besides, it has been well documented that second generation immigrants from Maghreb face the highest penalty on the French labor market among the different immigrant groups (see Algan et al, 2010, and Duget et al., 2010). Recent audit studies show that this labor market penalty is partly driven by pure cultural discrimination (Adida et al., 2010). We thus expect the transmission of Arabic names to raise an important trade-off between the desire to perpetuate one's own culture and the associated economic penalty inflicted to their offsprings with what can be perceived as a lack of integration with the dominant culture.

From this perspective, our key variable of interest is the choice of the first name instead of the last name since the latter one is not manipulable by the parent. Obviously the last name is also a marker of cultural identity. Besides, the last name can be correlated with productive characteristics that affect the economic penalty associated with culturally distinctive first name. Thus in the robustness analysis, we control for the nationality at birth of the parents as a measure of the origin of the last name and of ethnic backgrounds of individuals (color of skin, language...). This allows us to isolate, for a given last name, the specific signal associated to the transmission of a culturally distinctive first names.

In the last part of our paper, we provide several quantification and simulation exercises. We try in particular to evaluate the effects of different counterfactual experiments, which might be relevant in policy terms. Based on our structural model, we build various scenarios to assess the dynamic interplay between labor market policies, public housing policies and cultural transmission. For instance, we look at the equilibrium impact of lowering the extent of discrimination on the labor market for arabic-named individuals when naming patterns are endogenous.

In terms of results, we show that economic factors deeply shape the individual decision of cultural transmission. If the vertical channel associated with the culture of origin of the parents plays a key role in the cultural transmission, in line with the previous literature (Fryer and Levitt, 2004), we also find that parents do take into account the economic costs or advantages associated with their cultural trait in the naming decision of their babies. In particular, the unemployment penalty associated to Arabic name holders (which on the French labor market is around 5 percentage points) significantly reduces the probability for parents to give such names. The magnitude of the effect is also quite sizeable. Our estimates imply that, in absence of any unemployment penalty for Arabic name holders, the annual number of babies born with an Arabic name would increase by almost 60 percent, from the observed 201,895 births to a counterfactual 322,159 births.

Our empirical analysis is based on the French Labor Force Survey (LFS henceforth) from 2003 to 2007. The LFS combines three main features that provide a unique opportunity to estimate the various channels of cultural transmission. First, the LFS is a representative rolling panel at the household level that reports the first names of *all* the household members, including children,

and report detailed information on socio-economic characteristics. This unique information makes it possible to measure the vertical transmission channel. Second we can get from the LFS various proxies for the economic penalty associated with Arabic name as perceived by the parents. Third we can use residential neighbors as the source of information for the economic and horizontal channels since the data collection is based on close neighbors. The sampling unit in the LFS consists of groups (which we refer to as blocks) of adjacent households (on average 20 of them). In each wave, all the households and all members of the households, within the same block are interviewed. The LFS thus provides detailed information on both first names and socio economic characteristics of all the other individuals living in the block. This as a considerable data-related improvement upon the existing studies, which mostly look at neighborhood effects using spatial information from census data, and therefore have to rely on large areas (counties for instance) to define neighbors. Since close neighbors are likely to have a much stronger influence than distant neighbors on the individual decisions (a statement that we are able to quantify properly in our empirical analysis), our paper provides more accurate estimates of the horizontal and economic channels in the transmission of naming patterns.

A salient issue in the identification of the channels affecting the cultural transmission relates to the self-sorting of parents into different economic and neighborhood environments, leading to biased estimates of factors in the naming decision. In particular, two main biases arise from the identification of the economic and the horizontal channels. Regarding the estimation of the horizontal channel, a potential bias comes from the self-sorting of parents into given areas. The spatial clustering of Arabic names could thus be potentially driven by unobserved characteristics of the neighborhood. The same issue is raised by the identification of the economic channel. Parents attached to the transmission of the Arabic type to their offspring should, everything else equal, prefer occupations where the economic cost is lower. This potentially leads to a negative correlation between the economic cost and the unobserved heterogeneity in taste for Arabic names. We can alleviate this concern by focusing on the information that parents retrieve from their neighbors rather than from their own case. Thus our main source of identification is based on residential neighbors within a housing block for measuring both i) the peer-pressure for the horizontal transmission channel and , ii) a source of information for the perceived expected economic cost associated with a name. Yet, since individuals tend to self-segregate, e.g most households choose their location, our estimates of the horizontal and economic channel could be biased by spatial residential sorting.

Our identification strategy to address this bias is to focus on cultural transmission for households living in the French social housing sector. Due to an ideology deeply rooted in the French political system, and built into law, the government allocates state-planned moderate cost rental apartments to households without concern for their cultural background, mixing people indiscriminately. Furthermore, individuals rarely move, as the rents are much lower than market rates. We confirm with a variety of tests that spatial allocation inside the public housing market can be considered to a large extent as quasi-random. In particular, the observed residential allocation of household within social housing blocks is similar to a random distribution generated by Monte

Carlo simulation. Besides we show that while households moving into a new neighborhood tend to self-segregate in the unconstrained private housing market, there is no such evidence in the public housing market. We also examine potential self-selection prior to the move and show that households that have refused an offer end up living in public housing blocks that display the same characteristic, especially in terms of ethnic diversity, as those who accepted their first offer.

Our paper follows three main strands of related research. The first strand deals with the transmission of cultural values and the formation of identity (Akerlof and Kranton, 2000). Bisin and Verdier (2001) provides a seminal model of cultural transmission distinguishing vertical transmission inside the family from parents and oblique or horizontal transmission due to social interactions. Tabellini (2008) and Guiso et al. (2008) model the interactions between norms and economic incentives in the intergenerational transmission of values like trust. But those theoretical models have received little empirical support so far due to the lack of quantitative data on intergenerational transmission of culture. The closest paper to ours is Bisin and al. (2004) who use inter-religion marriages from the GSS to estimate a structural model of transmission of religious values. We distinguish from this literature in important dimensions. First, the naming decision provides an ideal experiment for estimating the determinants of cultural transmission. As emphasized by Lieberson (2000) the naming pattern is a "pure" expression of the parents' choice in the formation of cultural and social identity, since they are free and available to all parents without any material constraints. Alternative attributes of the cultural identity, such as wearing clothes with specific brands or religious values, are in contrast a costly investment that impose material constraints on the identity formation. This is obviously the case for cultural attributes associated with brands of consumption goods that could be manipulated by producers. But this is also true in the context of transmission of cultural values such as religious faith that could be constrained by the availability of churches, mosques or any other worship places in the neighborhood. Besides, naming practices provide an objective quantitative measure of cultural identity. Previous papers have used social surveys on attitudes and transmission of values (Bisin et al., 2010; Bisin et al., 2004), while the act of name-giving is a concrete cultural and social act. Second, our paper provides a structural model encompassing a new channel, associated to economic incentives, to the traditional horizontal and vertical channels discussed in the literature on cultural transmission. Third, previous papers on cultural transmission do not deal with endogenous spatial sorting to isolate the causal impact of the neighborhood (Bisin et al, 2010).

The second strand of the literature focuses more precisely on the determinants of naming patterns. Sociology was the first social science to analyze these issues. Lieberson and Mikelson (1995) study unique first names created by African American parents, and show that if such names reflect well the African origin of the parents (vertical transmission; e.g. root of the name), they are also to a large extent influenced by American norms in naming patterns (oblique transmission; e.g. sounding of the name). More recently, economists have sought to explain naming pattern, controlling more specifically for parents socio-economic background. Fryer and Levitt (2004) provide evidence that names given to children are an expression of cultural identity: they argue that the sudden rise

in the choice of distinctively Black names in the 1970's in the USA was a way to reinforce Black identity in the midst of the Black power movement, in line with the predictions of an identity model [1]. Our paper differentiates from Fryer and Levitt (2004) by identifying the interplay, and potential conflict, between economic and cultural incentives in the inter-generational transmission of culture. Fryer and Levitt (2004) find no conditional correlation between having a distinctively Black name and later life outcomes; by contrast, we find a pervasive unemployment penalty attached to Arabic first names and this specific feature of the French labor market enables us to analyze how economic factors affect the individual decision to transmit a naming pattern.

Head and Mayer (2008) measure the respective influence of socio-economic distance and geographical distance for the differences in naming patterns that emerge across French regions. While they are able to analyze nearly half a century of naming decisions, the lack of individual-level data makes it impossible for them to investigate the vertical and horizontal transmission channels in a sufficiently detailed way. Goldin and Shim (2004), focus on women's choice to keep or change their surname after wedding or child birth. Here, social determinants are examined (e.g religion, in laws) along with economic determinants (all the characteristics affecting her position/career achievement, such as education or age). They find that a woman will be more likely to keep her maiden's name if she has "made a name" in her career, highlighting the importance of economic background. We distinguish from this paper by focusing on the interplay between norms and economic incentives in the intergenerational transmission of first names.

The third related literature focuses on the effect of culture on economic outcomes. There is growing evidence that cultural values could shape individual economic behavior and affect aggregate economic outcomes (see Fernandez, 2008 for a recent survey). But we still have scant information on how the economic environment and economic incentives could shape cultural values and change their transmission pattern. Our contribution to this literature is that we are—to the best of our knowledge—the first to estimate how economic incentives affect individual decisions in the transmission of cultural values.

The remainder of the paper is as follows. Section 2 provides a detailed description of the data we use. Section 3 presents our theoretical model of naming decision, that will guide the estimated equation. This section also documents our identification strategy based on exogenous residential allocation of households in the public housing sector. Section 4 presents the benchmark estimation results. Section 5 quantifies the short-run effect of the vertical, horizontal and economic channels on cultural transmission. We also provide an alternative quantification by looking at the dynamic and long-run steady state effects of the channels. We run several quantification and simulation exercises based on our model of naming decision, fed with structural parameters estimated from the econometric work.

[1] They follow the Akerlof and Kranton (2000) identity model, where individuals gain a utility benefit when they behave according to what is "prescribed" by their type.

2 Data

2.1 The French Labor Force Survey

Our empirical analysis is based on the French Labor Force Survey (LFS henceforth) from 2003 to 2007. The LFS is a representative survey of the French population, stratified at levels of around 3500 residential blocks per year, each block being defined as an average of 20 *adjacent* households. The LFS is a rolling panel of 6 quarters and all the households within a given block are interviewed every quarter. All the household members aged above 15 year old are interviewed and they report information on their first names and their socio-economic characteristics, including the employment status (unemployed, inactive and employed), the hourly wage and the occupation. The occupation variable covers seven broad categories: farmer, craftsman, unskilled blue-collar, skilled blue-collar, clerk, intermediate, and executive. But the LFS also provides a more detailed classification of 29 occupations within those categories depending on the sector and infra-skill level of the occupation.

In addition, the survey records the first names of all the children in the household who are aged below 15 years old. This information allows us to estimate the economic and cultural factors in the naming pattern at a detailed individual level. Second, the LFS makes it possible to identify the determinants of the vertical transmission of first names, since we have information on the first names of both parents and children and we observe the choice of the first-name given to a child at birth. Third, the LFS provides a unique opportunity to understand the role of horizontal factors in the transmission of names since the data collection is based on (very) close neighbors. Given that the sampling unit in the LFS consists of groups of adjacent households, and that all the members of the households, within the same block are interviewed, we get detailed information on all the other individuals living in the close neighborhood. Another important characteristic of the LFS is to distinguish between the public and the private housing sector. As discussed below, our identification strategy will be based on residential allocation of households in the public housing sector. We thus report henceforth information both on the total sample and on the sample of public housing residential blocks.

The time span of the rolling panel is too short (6 quarters) to exploit the panel variation in the socioeconomic composition within residential blocks. We thus keep one observation per member of the household.[2] Table 1 reports the main descriptive statistics of the full database when we use the previous selection criterion. Our total sample is made up of 10'541 blocks, with 1'535 blocks belonging to the state-owned housing market. Each block consists on average of 18.31 adjacent households, each household being composed by around 3.31 members (babies, children and adults included). Overall, the total sample includes 425'210 individuals, among whom 69'458 are living in public housing.

[2] In general this observation corresponds to the first wave of interview of the block. If a baby is born in the subsequent waves of interview, we explain the naming decision by the socio-economic characteristic of the household and of the block that prevailed at the time of the first interview. We allow for a gap up to one year (4 quarters) between the explained outcome, e.g the choice of a baby's name, and the explanatory variables.

Table 1: Descriptive statistics of the residential blocks

	Total sample	Public housing
Number of blocks	10'541	1'535
Number of blocks by department	174.35	45.12
Average number of households per block	18.13	17.99
Average number of members per household	3.31	3.70
Average number of children per household (aged below 15 years old)	2.19	2.40
Total number of households	173'154	26'749
Total number of individuals	425'210	69'458

2.2 Sample of babies' names

Our main variable of interest is the individuals' name type and the cultural background that is associated to it. In particular, we focus on the transmission of Arabic names, as opposed to non-Arabic names, in the French society. As explained in the introduction, this focus is motivated on two main grounds. First, Arabic names are associated with the most important population of immigrants in France from Maghreb and Sub-Saharan Africa and to a lesser extent from the Middle East (mostly from Turkey), in the aftermath of the decolonization initiated in the 1960s. According to INSEE, people with Maghreb origins (i.e. Algeria, Morocco, Tunisia) represent 29.7 percent of migrants from first and second generations in France in 2008; this corresponds to 3.5 millions of individuals (1.6 million for the first generation and 1.9 million for the second generation) out of a total French population of 64.3 millions. Second, the Arabic names capture a cultural heritage that is potentially the most distinctive from the "locals", sometimes called *Français de Souche*, that is native French whose parents were also born in France. They are to a large extent a signal of the Muslim religious affiliation since most of those names come from the Qu'ran, and the transmission of first names associated with the Qu'ran is a natural practice for religious people. They are also associated in the French history with a hatred decolonization process such as the independence war in Algeria. In our data we identify the Arabic names by using the classification of Jouniaux (2001).

In addition to the name type, we have information on the country of birth and the nationality at birth of the respondents. Since we will focus in part of our analysis on naming decisions among first/second generation of migrants, we group together the country of origin and the nationality at birth in a corresponding category of Maghreb origin.

Table 2 displays the descriptive statistics of the sample of newborn babies. We observe 3'451 babies for whom we have all the relevant information on the parents' and blocks' characteristics. 3'216 babies (90.8%) receive a non-Arabic names. Among them, 1'879 babies (58%) are given traditional names, that is names that were already given in France in the early twentieth century.[3] Naming patterns that "sound" *Français de Souche* are thus still the most popular in the French society. Those traditional names are generally associated with Christian saint names, or names

[3]To identify those, we use INSEE's national database called "fichier des prénoms".

deeply ingrained in the French culture like Leo for boys or Manon for girls.

Table 2: Transmission of the name type across samples

	Babies with:	
	non-Arabic name	Arabic name
New Born (full sample):		
Parents with non-Arabic name	2,982 (489)	80 (28)
Parents with Arabic name	234 (95)	245(132)
0-3 years old (2nd/3rd generation):		
Parents with non-Arabic name	416 (183)	111 (47)
Parents with Arabic name	658 (317)	789(461)

Note: This table shows the full number of babies born with the two name types and allocates them according to the name type of their parents. The top panel gives figure for the whole sample of babies born within the year. The bottom panel considers babies aged 0 to 3 at time of survey, born from at least one parent or grandparent with nationality from Maghreb. In parentheses, the sample of babies in the public housing sector.

Among parents with Arabic names, the naming decision is rather balanced since 51.1 percent of those parents give an Arabic name to their offspring. Thus around one half of the parents from an Arabic cultural background transmit to their children a first name that sounds more traditional or more neutral relative to the French culture. In this latter case, they rarely choose saint names, but choose instead names that are culturally less distinctive. In particular, the two non-Arabic first names that are the more frequently selected are Adam or Yanis for boys, and Ines or Sarah for girls, names that seem to be attached to different cultures and are also given by the *Français de Souche* group of parents.

In contrast, among parents with non-Arabic names, the adoption of Arabic names is marginal, with a frequency of transmission of 2.8%. The top Arabic names given by those parents are Louna for girls and Rayan for boys. Those first names are rather neutral, they are hardly selected by parents with Arabic names.

Since we observe very little adoption of Arabic names by parents with non-Arabic names, we also look at the pure transmission decision of giving an Arabic name when it is part of the original culture. We focus on households where at least one parent or grandparent is a national from Maghreb. The babies are thus born in France, but the parents (babies of second generation) or the grandparents (babies of third generation) were born in Maghreb. Since this selection criterion would leave us with a too small sample of babies, specially in the public housing sector, we consider children between 0 and 3 years old instead of just newborn babies to carry out this analysis. While the use of this expanded sample of children could be a source of statistical noise in our econometric analysis, Section 4 shows that this problem is of little concern in practice.[4]

[4]Indeed, in our econometric analysis, we relate naming patterns to various observable characteristics of the residential block where the parents live. For newborn children the current block is the relevant one but this is not necessarily the case for children above the age of one. They may be born in a different neighborhood than the one we observe in our data (where each block is observed only during 6 semesters).

2.3 Employment penalty associated with Arabic names

The goal of this section is to document the extent of the economic penalty attached to Arabic first names across occupations. It should be clear that this investigation is only a first step for us, our final goal being to analyze in Section 4 how this economic penalty affect the individual decision to transmit a name type. We therefore keep our estimation method voluntarily simple in this first step. It is however interesting to note that our estimates of the penalty attached to Arabic first names are very much in line with the existing findings based on more elaborate econometric methods.[5] Another important point is that an accurate estimate of discrimination associated with an Arabic name—beyond its intrinsic complexity to obtain—is in fact not crucial for the purpose of this paper. Indeed , as will become clear below, our identification strategy is not based on the true penalty but rather exploits the *subjective* expected penalty that is perceived by parents after retrieving information from their neighbors (see equation 4 below).

Table 3 displays the unconditional unemployment rate associated with the 10 most popular non-Arabic and Arabic names. The data derives from the population aged 25-55 in the LFS 2003-2007. The unemployment rate of men with popular Arabic names is between four to eight times as high as the unemployment rate of men with popular non-Arabic names. A striking example is given by men named Abdelkader, whose unemployment rate reaches 37 percent against 5 percent for individuals named Philippe.[6] The unemployment gap is even more pronounced among the female population. Women named Fatma (ranked 6th in the list of the most popular Arabic names for women) have an average unemployment rate of 42 percent, against 10 percent for women named Sandrine or Patricia (who have the highest unemployment rate among the most popular non-Arabic names). The cross-name heterogeneity is also much larger for Arabic names but this is probably driven by small sample issues.

Table 4 breaks down the unemployment penalty associated with Arabic names by occupation. The average rate is also reported. For the sake of clarity (in this table only), we group together the 29 different occupations listed by INSEE into 7 main categories: farmer, craftsman, unskilled blue-collar, skilled blue-collar, clerk, intermediate, and executive. On average, Arabic name holders have an unemployment rate of 20 percent, around three times as high as the unemployment rate of non-Arabic name holders (7 percent). But this average comparison hides a lot of variance across

[5] In particular, Duguet et al. (2008) use pair auditing to test access to job interviews of individuals who share the same characteristics, except Arabic and non-Arabic names. They find that the probability to get interviewed is 7 percentage points lower for Arabic name holders in the French labor market, which is really close to our results. Adida et al. (2010) isolate the source of discrimination by identifying the effect of being Muslim on the French Labor Market. Using a large-scale survey on immigrants from Senegal, they are able to identify typical first names from the Muslim and the Christian parts of this population, which they report to be otherwise quite similar on all measurable aspects. The authors then run an audit survey with CVs identical in all dimensions, but with a different first name. The CVs would in particular have the same family name, for instance Diouf, a typical Senegalese one. But one CV would have a typical Muslim first name (for instance Khadija for women) and the other with a well-known Catholic first-name (Marie). Adida et al. (2010) find a statistically significant difference of 13 percentage points in the response's rates to job applications between the holders of Catholic first names and those with Muslim first names (which received 2.5 times less positive answers)

[6] Interestingly, emir Abdelkader was the military leader who led the struggle against the French colonial invasion of Algeria in the mid-19th century

Table 3: Unemployment rate by name

	Name type			
Rank of name	Non-Arabic		Arabic	
1	PHILIPPE	0.05	MOHAMED	0.19
2	ALAIN	0.05	SAID	0.30
3	CHRISTOPHE	0.07	RACHID	0.20
4	FREDERIC	0.07	ALI	0.20
5	PATRICK	0.07	ABDELKADER	0.37
6	MICHEL	0.05	KARIM	0.17
7	THIERRY	0.05	AHMED	0.18
8	PASCAL	0.05	MUSTAPHA	0.18
9	LAURENT	0.06	KAMEL	0.20
10	STEPHANE	0.08	FARID	0.29
1	NATHALIE	0.08	NADIA	0.18
2	SYLVIE	0.07	FATIMA	0.26
3	ISABELLE	0.08	MALIKA	0.24
4	CATHERINE	0.07	AICHA	0.26
5	CHRISTINE	0.06	NAIMA	0.15
6	MARTINE	0.08	FATMA	0.42
7	VALERIE	0.08	KHADIJA	0.17
8	SANDRINE	0.10	RACHIDA	0.37
9	VERONIQUE	0.08	SAMIRA	0.31
10	PATRICIA	0.10	YAMINA	0.36

Notes: The sample covers the 4 years of employment survey we have access to (2003-2007). The statistics are for adults between 25 and 55 years old.

occupations. The unemployment rate of Arabic name holders among executives is only 7 percent and the unemployment gap with non-Arabic name holders falls to 3 points for this occupation. In contrast, the unemployment rate of Arabic name holders reaches 29 percent among the (unskilled) blue collars, which represents an unemployment gap of 14 percentage points with the non-Arabic name holders belonging to the same occupational category. This variation across occupations suggests to estimate the employment penalty associated with an Arabic name at a very detailed level among the various occupations registered in the French national statistics.

Table 4: Unemployment rates by Name type and Occupation

	Unemployment rate:	
	Arabic name	non-Arabic names
Executive	0.07	0.04
Intermediate	0.14	0.05
Clerk	0.20	0.09
Blue collar (skilled)	0.20	0.07
Blue collar (unskilled)	0.29	0.15
Craftman	0.15	0.04
Farmer	0.10	0.00
Total	0.20	0.07

Table 5 documents the *conditional* employment penalty by running a standard Mincer-type equation estimated on the LFS subsample of active persons aged between 25 and 55. The left-hand-side variable is the employment status, equal to 1 if the respondent is employed, and 0 if unemployed. Logit regressions are performed with marginal effects reported in the table. We consider a set of standard controls, including nationality at birth of the respondent and parent's respondent, individual characteristics (age, age squared, gender, marital status and number of children), educational, occupational, spatial and year fixed effects. Our variable of interest is *arabic name*, a binary variable coding for a first name from Arabic origins. Column (1) reports the employment penalty associated with an Arabic name, without controlling for the nationality at birth of the respondent and of the parents' respondent. Holding an arabic name decreases the probability to be employed by 10 percentage points and the effect is statistically significant at the 1 percent level. However, most of Arabic name holders being first or second generation migrants, the previous correlation captures both the discriminating impact on the labor market of foreign origins and of foreign names; while closely related, the latter dimension is manipulable by parents but the former is not. To isolate the specific penalty from a name that sounds culturally distinctive, we also control for other attributes of the country of origin. Column (2) includes a dummy variable equal to 1 if the nationality at birth of the respondent or of the parents' respondent is from Maghreb, and 0 otherwise. The estimated employment penalty associated with an Arabic name remains fairly high at 7 percentage points and remains highly statistically significant. The estimated employment penalty associated with an Arabic Name is of the same order of magnitude as the one associated with having a Maghreb-related nationality, suggesting that a specific employment penalty is attached

with the first name.[7] In the next two columns, the sample is restricted to individuals living in the public housing sector (on which our main econometric analysis will be based). Column (3) shows that the conditional employment penalty is identical for this subsample and is robust to the inclusion, in Column (4), of a variable coding for the number of children *with* Arabic name in the household. This last variable is likely to be correlated to a bundle of *unobservable* characteristics related to the degree of attachment of the individual to the Arabic culture (e.g. religiosity) that may simultaneously influence the penalty. Column (5) estimates the conditional employment penalty for each broad occupational category. The reference category is executives.

How big is the implied loss in lifetime expected income associated with an Arabic name? A simple"back of the envelope" calculation suggests that the loss is substantial, reaching 24,382 euros, which is roughly equivalent to the loss of one year of average income. [8]

Table 5: The penalty of an Arabic name

	(1)	(2)	(3)	(4)	(5)
Dep.Var:		emp. / unemp. status			
arabic name	-0.10^a	-0.07^a	-0.07^a	-0.06^a	0.00
man	0.01^a	0.01^a	0.03^a	0.03^a	0.01^a
age	0.01^a	0.01^a	0.01^a	0.01^a	0.01^a
age squared	-0.00^a	-0.00^a	-0.00^a	-0.00^a	-0.00^a
nationality from Maghreb/Middle-East		-0.06^a	-0.07^a	-0.06^a	-0.06^a
count of kids with Arabic name				-0.01^b	
arabic × intermediate					-0.06^a
arabic × clerk					-0.08^a
arabic × blue collar (skilled)					-0.09^a
arabic × blue collar (unskilled)					-0.09^a
arabic × craftsman					-0.08^a
arabic × farmer					-0.04
Observations	148582	148582	90693	90693	148582
R^2	0.041	0.042	0.047	0.047	0.049

Note: Column (5) has executives as the baseline occupation group. All regressions include dummies for education level, occupation group, *département* of residence, years, as well as number of children, and marital status. The sample includes active persons aged between 25 and 55. The unconditional unemployment rate in this sample is 8%.

[7]This might reflect the absence of clear morphological markers of ethnicity (e.g. skin color, size) for individuals with origins from Maghreb and living in France. Hence the first name conveys meaningful information on the ethnical background.

[8]From a simple "back of the envelope" calculation, we can assess the income loss associated with the employment penalty of Arabic names holders. We use Breuil-Grenier's (2001) detailed estimates of the income variation induced by a transition from employment to unemployment on the French labor market. The author takes into account the level of all the social benefits associated with an unemployment spell and the eligibility criteria. She finds an average income loss of 50 percent. The average monthly income of an employed individual in the 2003-2007 LFS is 1,488 euros. Unemployed people thus earn on average 744 euros per month. From Table 5 - Column (2), we know that every period, the conditional unemployment gap of Arabic name holders is 7 percentage points relative to non-Arabic name holder. They thus have an expected income loss of $0.07 \times 744 = 52.1$ euros per month. Since the average participation to the active population is 39 years in France, this means that the total income loss of typical Arabic name holder during their active life reaches $39 \times 12 \times 52.1 = 24,382$ euros, which is roughly equivalent to the loss of one year of average income (the average monthly income is of 1800 euros during the period at stake)

3 Model and Identification of Naming Decision

3.1 A simple model of baby name choice

In this section we build a random utility discrete choice model of baby naming decision. Our framework is rich enough to highlight the underlying estimation issues while remaining sufficiently tractable for structural estimation. The parental decision under scrutiny is binary and relates to the cultural type attached to the baby's first name. The utility for a household i, living in residential block k, derived from choosing a given name type for its baby born in year t is defined as $U_{ik,t}(1)$ if the name is Arabic and $U_{ik,t}(0)$ otherwise,

$$U_{ik,t}(\text{Baby}) \equiv V_{ik,t}(\text{Baby}) + \epsilon_{ik,t}(\text{Baby}), \tag{1}$$

where $\text{Baby} \in \{0, 1\}$ denotes alternatives, $V_{ik,t}(\text{Baby})$ is the observed part of utility and $\epsilon_{ik,t}(\text{Baby})$ is the unobserved parental-specific random shock across alternatives.

In such a discrete choice setting, only differences in utility over alternatives can be identified from the data. The econometrician observes a parental choice $\text{Baby}_{ik,t} = 1$ if and only if $\Delta U_{ik,t} \equiv U_{ik,t}(1) - U_{ik,t}(0) \geq 0$. Let us denote the difference in the observed part of utility as $\Delta V_{ik,t} \equiv V_{ik,t}(1) - V_{ik,t}(0)$, and the difference in unobserved utility as $\varepsilon_{ik,t} \equiv \epsilon_{ik,t}(1) - \epsilon_{ik,t}(0)$, such that

$$\begin{aligned} \Delta U_{ik,t} &= \Delta V_{ik,t} + \varepsilon_{ik,t} \\ &= \alpha_0 + \alpha_1 \underbrace{\text{Parents}_i}_{\text{Vertical}_i} + \alpha_2 \underbrace{\mathbb{E}\left[\frac{\sum_{j \in k, j \neq i} \text{Baby}_{jk,t}}{\mathcal{N}_{k,t}}\right]}_{\text{Horizontal}_{k,t}} + \alpha_3 \underbrace{\mathbb{E}[\mathcal{C}_{ik}]}_{\text{Economic Cost}_{ik}} + \varepsilon_{ik,t}, \end{aligned} \tag{2}$$

where $\Delta V_{ik,t}$ is specified as a three part linear function, which we label "Vertical", "Horizontal", and "Economic cost" channels of influence. Parents_i is a parental characteristic equal to one when the name of one of the two parents is Arabic and zero otherwise (with alternative definitions investigated in Section 4), $\text{Baby}_{jk,t}$ codes for choices of names among the $\mathcal{N}_{k,t}$ other babies born in t and living in residential block k and $\mathbb{E}[\mathcal{C}_{ik}]$ is the *perceived* economic penalty that parents i expect to be attached to their baby if they choose an Arabic name.

The first RHS component therefore corresponds to the desire by parents to transmit their own cultural type (measured by the coefficient α_1). Our specification of utility is flexible as it allows both for cultural transmission and cultural adoption. Transmission is the case where the names of parents and babies belong to the same cultural type. Adoption corresponds to the two other cases: e.g. parents with Arabic names that do not transmit their cultural type to their baby or parents with non-Arabic names adopting an Arabic name for their baby. Both patterns are observed in the data although the latter is less salient (see Table 2).

The second RHS component reflects social influence, i.e. the share of parents of newborn babies in residential block k expected to make the same choice as i, with intensity α_2 expected to be positive. In our data, the block k is small enough that household i is not negligible and this

creates the classical Manski (1993) reflection problem. Following Glaeser and Scheinkman (2000) we assume that parents i form their expectations on lagged decisions of neighbors[9]:

$$\mathbb{E}\left[\frac{\sum_{j\in k,j\neq i}\text{Baby}_{jk,t}}{\mathcal{N}_{k,t}}\right] \equiv \frac{\sum_{\tau=1}^{\Upsilon}\sum_{j\in k,j\neq i}\text{Baby}_{jk,t-\tau}}{\sum_{\tau=1}^{\Upsilon}\mathcal{N}_{k,t-\tau}}, \tag{3}$$

that is they expect the current choices of neighbors to be on average similar to the ones taken since year $t-\Upsilon$ (we will take $\Upsilon = 10$ in our application).

The third RHS component relates to economic incentives: Presumably, the higher the expected penalty is, the less parents want to attach an Arabic cultural type to the name of their babies. The comparison of the coefficients α_1 and α_3 reflects the parental tradeoff between their own attachment to a particular cultural type and their altruistic concern towards the future economic performance of their babies. The *perceived* expected penalty, $\mathbb{E}[\mathcal{C}_{ik}]$, is sensitive to the parental information set and to a wide set of observed and unobserved parental characteristics influencing the future spatial and social mobility of the baby. Our identification strategy exploits the fact that part of this information set is based on information on the labor market that households retrieve from social interactions and communication with their neighbors. The idea is that parents surrounded by neighbors working in occupations with high levels of penalty tend to update upwards their beliefs on the extent of the penalty. Formally, the *perceived* expected penalty is broken down into a block-specific informational component and an unobserved parent-specific residual component:

$$\mathbb{E}[\mathcal{C}_{ik}] = \sum_{l\in\mathcal{O}}\omega_{lk}\times\hat{\gamma}_l + u_i, \tag{4}$$

where $\mathcal{O}=$ is the set of occupations, ω_{lk} is the share of neighbors working in occupation l in block k, $\hat{\gamma}_l$ is a proxy for the labor market penalty observed in occupation l and u_i is the residual parent-specific part. In the remainder of the paper, $\sum_{l\in\mathcal{O}}\omega_{lk}\times\hat{\gamma}_l$ is labeled as the *block information on penalty*.

Combining (2), (3) and (4), utility becomes

$$\begin{aligned}\Delta U_{ik,t} &= \Delta\mathcal{V}_{ik,t} + \delta_{ik,t} \\ &\equiv \alpha_0 + \alpha_1\text{Parents}_i + \alpha_2\frac{\sum_{\tau=1}^{\Upsilon}\sum_{j\in k,j\neq i}\text{Baby}_{jk,t-\tau}}{\sum_{\tau=1}^{\Upsilon}\mathcal{N}_{k,t-\tau}} + \alpha_3\sum_{l\in\mathcal{O}}\omega_{lk,t}\times\hat{\gamma}_l + \delta_{ik,t},\end{aligned} \tag{5}$$

where $\Delta\mathcal{V}_{ik,t}$ is the observable utility and $\delta_{ik,t}\equiv\alpha_3 u_i + \varepsilon_{ik,t}$ is the new error term.

In this equation, our source of identification is based on neighbors within a residential block. Neighbors serve both as: i) A source of peer-pressure for the horizontal transmission channel, and ii) A source of information for the perceived expected penalty associated with a name. Since individuals tend to self-segregate, e.g most households choose their location, our estimates of equation

[9] An alternative way to mitigate the problem could be through an IV strategy, where natural instruments of parental expectations are some observable characteristics of the neighbors (the instruments should not influence the choice of household i *directly* but through the choice of surrounding parents j).

could be biased by endogenous residential sorting. To address this bias, we restrict our estimations to a subsample of randomly allocated households, namely those living in the public housing sector. In the next section, we document step by step the issues raised by the identification of equation (5). In Section 3.3 we document and test the fact that the residential allocation within the French public housing sector is quasi-random and exogenous with respect to ethnic characteristics.

3.2 Estimation issues

The horizontal transmission channel raises several estimation issues well-known in the social interaction literature (for a recent survey of discrete choice models with social interactions, see Blume, Brock, Durlauf and Ioannides, 2010). Indeed, in equation (5), the realizations of $\text{Baby}_{jk,t-\tau}$ depend on $\Delta U_{jk,t-\tau}$. *Spatial sorting* might lead to a non-zero correlation between $\delta_{ik,t}$ and $\delta_{jk,t-\tau}$ for households i and j belonging to the same residential block k. This would create a correlation between $\text{Baby}_{jk,t-\tau}$ and the error term in (5), $\delta_{ik,t}$, potentially capturing unobservable taste shocks for the considered cultural type common to households i and j. For example, it is clear that the degree of religiosity of the household, which is unobserved by the econometrician, affects positively the choice of an Arabic Name for the baby; moreover religious people tend to live in same residential areas (e.g. close to a Mosque or to Halal shops). This example makes it clear that spatial clustering of Arabic names is not only driven by horizontal transmission but is also potentially partly driven by unobserved characteristics of the area. Our estimates could thus be biased by the endogenous spatial sorting of households. To get rid of this bias, we identify the coefficient α_2 on the subsample of households that are randomly allocated across the different public housing blocks within *départements*.[10] This *within* identification strategy calls for including *département* fixed effects in all specifications.

The coefficient α_3 associated with the economic cost of a name type is also potentially ill-estimated due to self-selection into occupations and locations by parents. The concern is that religious (muslim) parents, attached to the transmission of the Arabic type to their offspring, tend to work in occupations with low discrimination, and are located in residential blocks with religious neighbors working in non-discriminating occupations. We address this issue first by controlling for parental occupation and education fixed effects. Although parental occupation is not a random choice, the inclusion of fixed effects for the parental occupation captures all time-invariant co-determinants of the parental occupation choice and the naming pattern. Second, rather than using the parental occupation as a source of information on the perceived expected penalty, we use the block-specific informational component of the penalty. Thus the key issue is now that of the exogeneity of the composition of occupations within the residential block. We identify the coefficient α_3 by restricting once again our estimates to the subsample of randomly and exogenously allocated households living in the public housing sector within a given *département*.

As discussed earlier, our goal is not to identify accurately the discrimination associated with an Arabic name. In our empirical application we proxy $\hat{\gamma}_l$, the labor market penalty observed in

[10] Metropolitan France is divided into 95 administrative areas, called *départements*.

occupation l, as the unconditional unemployment gap between Arabic and non-Arabic name holders in occupation l. The fact that $\hat{\gamma}_l$ is an imperfect measure of the true penalty in occupation l could lead to an attenuation bias in the estimation of α_3. We will consequently consider an alternative and more accurate measure of $\hat{\gamma}_l$ based on the *conditional* unemployment gap as computed in Table 5. The comparison of the estimates obtained with each of the measure (Columns 3 and 6 in Tables 8 and 9) confirms a very light, but statistically insignificant, increase in α_3 when estimated with the second measure, showing that this attenuation bias is unlikely to be a first-order issue.

3.3 Identification with exogenous spatial residential allocation in public housing

As discussed above, the estimation of the horizontal and economic channels of cultural transmission raises the issue of endogenous residential sorting of households. Individuals might tend to self-segregate if they prefer to live close to neighborhoods with whom they share common characteristics, in particular people from the same ethnic and socio-economic backgrounds. We claim that the French public housing provides a source of exogenous residential allocation to mitigate this bias. Since public housing is administrated at the *département* level, we investigate whether households are allocated randomly across the different public housings within a given *département*.

3.3.1 Formal allocation process in public housings

We start by documenting the actual process of allocation of households across public housing dwellings. More details on the institutional and legal aspects are provided in Appendix A.1. The main eligibility requirements for admittance into the public housing sector are to be legally living in France (as a French citizen or migrant with a valid residence permit) and to be living under a certain threshold of income per unit of consumption. This income ceiling is usually rather high: in 2009, this threshold was between 36'748 and 50'999 Euros per year for a four-person family, depending on the département of residence. Jacquot (2007) estimates that around two thirds of households living in Metropolitan France could apply for a public housing unit. Moreover the rents are considerably lower in public housing than in private housing. As a result, there is a strong excess demand for public housing. Just for the case of Paris, there were 121'937 ongoing applications, to be compared to 12'500 public housing units allocated over the year 2010. Due to those stringent constraints, other eligibility criteria are taken into account : the family situation and household size (to ensure a suitable match with the characteristics of vacant dwellings), as well as the emergency of the application.[11]

The selection committees in charge of allocating households to vacant public housing dwellings are held at the *département* level. For each vacant housing unit, at least three households must be considered by the committee members, who finally decide which household will be allocated to which housing unit according to the eligibility and priority criteria detailed above. The application form contains very limited information about the ethnicity of the applicant: he or she only needs to

[11] Five priority criteria—none related the nationality—are defined by law to make sure that vacant housing will first be distributed to households with obvious social difficulties (see Appendix).

inform about his or her nationality, which is limited to three possible categories (French, European Union, or non European Union). Legally, applicants can refuse up to three offers but in practice they rarely do, given the large opportunity cost of declining an offer. This makes it unlikely that the selected households could be really picky about the characteristics of their neighborhood. It is also formally possible to indicate a precise neighborhood within the département in the application form, but in practice, very few applicants (6.6 percent) do provide this information. For example in Paris, where public housing blocks are scattered around all the areas, more than half of the 121'937 applicants (52.9 percent) did not mention any particular area at all, probably due to the fear of being rejected on this ground. Residential mobility within the public housing sector is very low, due to the current strong shortage in the supply of public housing dwellings. People who move within the public housing sector are people who moved for larger space following an increase in their household size (only 12 percent of the public housing dwellings have more than three rooms). Hence, residential self-sorting, especially on ethnic characteristics, is not a common practice within public housing (Simon, 2003).

In short, the public housing market is very tight, and highly regulated. This implies that households have very limited control over the time when they will be assigned to a public housing dwelling and the precise place within a département where they will be located.

3.3.2 Statistical tests of exogenous residential allocation in public housings

We now provide more formal statistical tests to show that the spatial allocation of households across public housing residential blocks within a département is random.

Table 6 provides a first approach where, for various observable household characteristics, we test for the difference in means between residential blocks. We regress, for each *département* taken separately, each observable on a battery of fixed effects associated to the different residential blocks located in this *département*. Those regressions are performed on the subsample of household heads who are living in public housing. The random allocation test consists in performing a standard F-test on the null hypothesis that the fixed effects are jointly not statistically different from zero. In the case of endogenous residential sorting in some public housing blocks, the fixed effects associated with those blocks should be statistically significantly correlated with the household characteristics, and the F-test will be rejected. We consider four different households' characteristics: the Arabic nature of the respondent's name, the nationality at birth (set to one if the nationality of the respondent or of the respondent's parent is from France, and zero otherwise), the occupation of the respondent (coded as a binary variable equal to one for blue collars), and the education level (measured by "No elementary school", this educational level being over-represented in public housing). Column (1) of Table 6 reports, for each observable, the share of département for which the F-test is not rejected. For the sake of comparison, we run, in Column (2), the same F-test on the full sample of household heads, including both those who live in the public and private housing sectors. In this case, as expected, endogenous residential sorting is much more salient.

The previous test of random spatial allocation, though simple, is not ideally suited to our

Table 6: F-Test of Residential Sorting

% departments without residential sorting relative to households' characteristics

	Public Housing	Total Sample
Household's characteristics		
Arabic names	86.53	67.85
French Nationality at Birth	83.01	66.66
Occupation: blue collar	96.72	67.02
Elementary education	98.36	64.89

Note: The table reports the share of *département* for which F-tests (at the 5 percent level) do not reject the null-hypothesis of a null correlation between observable characteristics and residential block fixed effects. The F-test are based on Logistic regression of household characteristic on public housing fixed effect within each *département*. The sample includes household heads aged over 15 years old.

empirical context where the spatial units under consideration are small. Indeed, in the LFS, the average residential block is composed of only 18.4 household heads. In this type of context, as firstly pointed out by Ellison and Glaeser (1997), any parametric test of spatial allocation/concentration (that assume independent location choices) will tend to *over-reject* the null hypothesis of random allocation.[12] We follow the literature and perform Monte Carlo simulations for generating artificial random allocations that we compare to the actual observed allocation. More precisely, in our second test, we pool the public housing population and reallocate it randomly, without replacement, across the different residential blocks of the corresponding *département*, maintaining unchanged the actual size of each block. After 10000 draws of Monte Carlo allocations, we compute, for each *département*, the average simulated random distribution of native French and Maghrebian shares across blocks and we perform two tests of equality with the actual spatial distribution - a simple t-test of equality of means of the distribution and a more-demanding Kolmogorov-Smirnov test. Table 7 reports the percentage of *départements* for which the actual and simulated distributions of ethnic shares across neighborhoods are similar, i.e. those for which we cannot reject the null hypothesis of equality of the distribution at the 5% level. The labels in the first column indicate the ethnic group under consideration, the second column reports the results of the t-test and the third column reports the Kolmogorov-Smirnov test results. In addition, we perform the same test on two other characteristics than ethnicity: whether the household head has achieved primary education or not, and whether he is a blue-collar worker or not. The results displayed in the last two lines of Table 7 also show

[12] As mentioned previously, the allocation of households across public housing blocks takes place at the *département* level. If the members of the public housing commission strictly follow the legal criteria and do not take into account the ethnic characteristics in the allocation process, we should find an uniform distribution of households of a given nationality across the various public housing residential blocks *only if the size of each block is large enough*. For the sake of illustration, let us assume that 10 percent of Maghrebians live in the public housing sector in Paris. We should find the same share of 10 percent of Maghrebians within each Parisian housing block if the allocation was truly exogenous with respect to ethnic characteristics only if we have a sufficiently large number of individuals within each housing block; otherwise we will observe pattern of spatial concentration in some blocks. The F-test propensity to *over-reject* the null hypothesis of random allocation is confirmed by the comparison of the results in tables 6 and 7 .

that the distribution of the population across public housing is close to random when it turns to education and socio-professionnal characteristics as well. All in all, those tests are in line with our identifying assumption that the allocation of households across the public housing blocks of a given locality is quasi-random.

Table 7: Monte-Carlo Test of Random Allocation within Public Housing

	t-tests	**K.S-test**
Native French	100%	98.61%
Maghrebi	100%	98.01%
Both groups	100%	97.22%
Occupation: blue collar	100%	98.18%
Elementary Education	100%	99.32%

Note: Comparison between the actual and simulated distributions by ethnic groups shares, education and occupations across public housing blocks: percentage of *départements* where equality is not rejected.

In Appendix A.3, we perform a variety of additional tests. First, we focus on movers and show that, in the public housing sector, movers do not self-select select new neighborhoods where their ethnic and socio-economic characteristics (education and occupation) is over-represented. Since self-selection could occur prior to the move, we also look at the characteristics of households that have refused a public housing dwelling offer. We show that they display the same characteristics as those who accepted their first offer. Thus even if households try to be choosy with respect to the ethnic and socio-economic composition of their neighborhoods, they cannot self-segregate in the public housing sector due to the allocation process and the tight supply constraints of dwellings. Appendix A.2 displays descriptive statistics in the public housing sector.

4 Econometric Results

In this empirical section we estimate the utility function (5) while addressing the identification issues discussed in Section 3.2. To proceed, we need to go from the utility function to an estimable discrete choice econometric model. A standard logistic distribution for the error term $\delta_{ik,t}$ is specified, with σ the scaling parameter and μ the location parameter of this distribution. One can then express in closed-form the probability of choosing an Arabic name – a formula that enables, in Section 5, to run counterfactuals without probabilities going out of bound:

$$\mathbb{P}_{ik,t}(\text{Baby} = 1) = 1/[1 + \exp(-\Delta\mathcal{V}_{ik,t}/\sigma)], \tag{6}$$

The observable utility differential $\Delta\mathcal{V}_{ik,t}$ is retrieved from the coefficients in (5) that can be readily estimated through a standard Logit - Maximum Likelihood procedure (up to the scaling parameter). The dependent variable is a binary variable coding for the Arabic origins of a baby's name; explanatory variables relate to various parental and neighbor characteristics. All specifica-

tions include parental occupation fixed effects, parental education fixed effects, and spatial fixed effects at the *département* level (see Section 3.2). Average marginal effects are reported in tables.

4.1 Babies from all origins

Table 8 reports the benchmark estimates for the sample of babies born over the 2003-2007 period and from all ethno-linguistic origins. The vertical transmission channel is represented by two binary variables coding for the type of parental first names and for the nationality of origins of the parents/grandparents. The former captures cultural transmission/adoption (as discussed in Section3.1) and the latter isolates the specific effect attached to 2nd/3rg generation of migrants from Maghreb. Horizontal transmission is measured by the *share of name type in the block* (defined by 3)) and our proxy for the economic cost of a name corresponds to the *block information on penalty* that is based on the unconditional unemployment rate differentials by occupation between Arabic and non-Arabic name holders, weighted by the shares of occupations at the block level (see equation (4) and our discussion at the end of Section (3.2)).

Column (1) reports the result for newborn babies in the full sample of private and public housings. All four coefficients of interest are statistically significant and have the expected sign. The t-stats of the two coefficients on vertical transmission are particularly high; this reflects the very stark contrast in the pattern of cultural transmission between parents of Arabic origins and parents without such origins, a feature of the data which has already been discussed in Section 2.2. The horizontal transmission channel is also significant at the one percent threshold. The coefficient associated to the block information on penalty is negative and statistically significant at the five percent level confirming that parents take into account the employment penalty of Arabic names in their own occupation when they transmit a first name: they are deterred from inflicting an economic cost to their offspring. We provide in Section 5 several quantifications for the order of magnitude of the effects.

In Column (2) we address the endogenous residential sorting issue by restricting the sample to newborn babies of households living in public housing. With one quarter of the original sample left, the number of observations drops dramatically. The coefficients keep the expected sign but they are now much more imprecisely estimated. To overcome this lack of data (see the discussion of table 2), we expand in Column (3) the sample to all children aged between 0 and 3 of households living in public housing, instead of restricting to newborn. To avoid any overlap, we adjust the age definition to measure the horizontal channel by considering babied aged between 4-10 in the housing block. The sample size in the public housing sector is now comparable to the one with the original full sample in Column (1). Column (3) shows that the coefficients keep the same magnitudes than the ones in Column (2) with a statistical significance now restored at the 1 percent level.

Columns (4) and (5) document interesting features of the horizontal channel. In column (4) we add the share of arabic names for kids under 10 in a larger area, either at the *sector* level, which consists of 6 adjacent housing blocks, or at the *département* level. Those two variables measure the spatial decay of the horizontal transmission channel by looking at wider geographic

unit. None of those two variables exhibit any influence, and the block-based horizontal estimate is unchanged. This points to the importance of studying those channels of transmission at a very fine-grained geographical level (namely a block of 20 adjacent households), which is impossible with data usually at hand in the social interaction literature. Column (5) includes the share of arabic names for older cohorts to identify the groups to which the parents refer to in their naming decision. Only the coefficient of the cohort of kids under 10 years old has a statistically significant impact in their naming decision. It make us confident when concluding that this variable captures horizontal influence and that, in our sample of public housing, our estimates are unlikely to be pervasively contaminated by endogenous residential sorting (otherwise the coefficients for older cohorts should also be non zero).

In Column (6) the *block information on penalty* is now based on the *conditional* unemployment rate differentials by occupation between Arabic and non-Arabic name holders as computed in the Mincer-like specifications of Table 5. The coefficient associated with the economic penalty in the naming decision remains statistically significant at the one percent level and of the same order of magnitude. The stability of this coefficient in spite of conditioning the penalty by a large set of observables confirms that the attenuation bias due to measurement errors on the true level of the unemployment penalty is unlikely to be a first-order issue in our estimates.

4.2 Babies from 2nd/3rd generations of immigrants from Maghreb

The previous estimates mixed up the decisions of cultural transmission/adoption by parents from various cultural background. We now estimate the different channels when we focus on pure transmission, that is by focusing on the determinants of naming decisions for babies who are second or third generation of migrants from Maghreb, e.g who are born in France while their parents or grand-parents are born in North Africa. Table 9 replicates the specification of Table 8 on this new sample. We first report the results for the sample of new born babies, in both the private and public housing sector (Column (1)). In order to eliminate the residential sorting bias the next specifications are based on the public housing sample. The sample of newborn babies from Maghrebi origins in the public housing sector (Column (2)) is very small (and even lower than in Column (2) of Table 8) and coefficients are imprecisely estimated. Column (3) displays the result for 2nd/3rd generations of children aged between 0 and 3, instead of restricting to newborn, increasing by a factor of five our sample (992 observations). Remarkably, the effect of the block information on penalty is around 3.5 times as large for parents from Maghreb origins as for the full sample of parents (Column (3) of Table 8)). In contrast the horizontal channel does not matter for this sample of parents, suggesting a polarization of the trade-off between the vertical and economic channels in the naming decisions. Columns (4) and (5) provide alternative specifications for measuring the effect of the horizontal channel by looking at the peer effect of older children or from larger geographic localities, without any significant change. Column (6) shows that the coefficient of the economic channel remains stable and statistically highly significant when we use the conditional unemployment penalty instead of the unconditional penalty for building the block information on

Table 8: The choice of an Arabic name - All Origins

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Var:	Arabic name for baby					
one of grandparents/parents has Maghreb-related nationality	0.07[a]	0.08[b]	0.09[a]	0.09[a]	0.09[a]	0.09[a]
	(0.01)	(0.03)	(0.01)	(0.01)	(0.01)	(0.01)
at least one parent same name type	0.12[a]	0.27[a]	0.23[a]	0.23[a]	0.23[a]	0.23[a]
	(0.01)	(0.03)	(0.01)	(0.01)	(0.01)	(0.01)
share of name type in block (aged 1-10)	0.05[a]	0.06				
	(0.02)	(0.06)				
block information on penalty	-0.43[b]	-1.14	-0.86[a]	-0.92[a]	-0.97[a]	
	(0.21)	(0.79)	(0.33)	(0.33)	(0.34)	
share of name type in block (aged 4-10)			0.09[a]	0.10[a]	0.07[a]	0.07[a]
			(0.02)	(0.02)	(0.02)	(0.02)
share of name type in sector (aged 4-10)				0.01		
				(0.03)		
share of name type in dept (aged 4-10)				-0.16		
				(0.13)		
share of name type in block (aged 11-25)					0.02	0.03
					(0.04)	(0.04)
share of name type in block (aged 26-49)					0.06	0.06[c]
					(0.04)	(0.04)
share of name type in block (aged 50+)					-0.04	-0.04
					(0.03)	(0.03)
block info. on penalty (Mincer-based)						-1.30[a]
						(0.36)
Only social housing	No	Yes	Yes	Yes	Yes	Yes
Age of babies	0	0	0-3	0-3	0-3	0-3
Observations	3541	660	3829	3811	3777	3777
Pseudo R^2	0.446	0.443	0.399	0.402	0.402	0.403
Average prob.	0.09	0.21	0.19	0.19	0.19	0.19

Note: logit estimates (average marginal effects). Standard errors in parentheses with [a], [b] and [c] respectively denoting significance at the 1%, 5% and 10%. All regressions include dummies for parental education level, occupation group, *département* of residence, and years.

penalty.

4.3 Additional Robustness checks

All robustness checks are provided for the subsample of children in social housing, aged between 0 and 3 and from Maghrebi origins (2nd/3rd generations). Due to the drop in sample size, those are the most demanding specifications.

In Table 10 the first column displays the benchmark estimates of column (3) in Table 9. In Column (2) we report the results for a weighted logit, where the individual representativeness weights reported in the LFS are applied. As discussed in Section 2, the labor force survey is stratified at the *département* level and representativeness is thus not guaranteed at the block level, our level of analysis. Our 992 children in column (1) therefore represent 618'314 children at the national level. Coefficients are only slightly affected.

In the next two columns we run placebo tests to rule out the possibility that our estimate of the economic cost could be driven by some residual statistical bias attached to endogenous residential sorting. We replicate our benchmark specification on a fake sample of parents/neighbors artificially reallocated to random occupations in Column (3) and to random residential blocks in Column (4). [13]. We see that in both cases the block information on penalty, that is based on neighbors occupations (see definition 4), drops and also looses its statistical significance. This makes us confident that our identification strategy, based on the sample of children living in public housing, gets rid of endogenous residential sorting in an efficient way. In columns (5) and (6), we test whether our effects depend on the gender of the children by splitting the sample in girls and boys respectively. We see that the horizontal transmission channel matters more for girls while the economic cost channel is more important for boys.

Columns (1) and (2) of Table 11 document the effects for couples where both parents have Maghreb origins and for mixed couples, respectively. The magnitude of the economic cost channel (though less significant) is larger for the latter. Columns (3) and (4) display separately the results for 2nd generation children only (i.e. babies with parents who migrated to France) and for 3rd generation children only (i.e. babies with parents who are born in France). Clearly, the economic cost channel is much larger for the latter. Our interpretation is that parents from Maghreb origins who are born in France are more exposed to information on discrimination. In the last two columns, we differentiate the effects within public housing between the normal residential blocks and the residential blocks from deprived areas that are officially classified as ZUS (*Zone Urbaine Sensible*). Here again we find that the magnitude of the economic cost is much larger in the latter confirming that residents from those areas are much more sensitive to discrimination.

[13] XXX Thierry, tu peux donner un peu plus de dtail ici stp. Combien de reallocation?

Table 9: The choice of an Arabic name - Migrants from Maghreb

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Var:	Arabic name for baby					
at least one parent same name type	0.33[a]	0.39[a]	0.36[a]	0.36[a]	0.35[a]	0.35[a]
	(0.05)	(0.09)	(0.04)	(0.04)	(0.04)	(0.04)
share of name type in block (aged 1-10)	0.05	0.04				
	(0.08)	(0.14)				
block information on penalty	-1.57	-5.01[c]	-2.95[a]	-3.17[a]	-3.32[a]	
	(1.25)	(2.74)	(1.07)	(1.06)	(1.03)	
share of name type in block (aged 4-10)			0.03	0.05	-0.01	-0.02
			(0.05)	(0.06)	(0.06)	(0.06)
share of name type in sector (aged 4-10)				-0.12		
				(0.11)		
share of name type in dept (aged 4-10)				0.10		
				(0.36)		
share of name type in block (aged 11-25)					0.05	0.07
					(0.10)	(0.10)
share of name type in block (aged 26-49)					0.09	0.11
					(0.10)	(0.10)
share of name type in block (aged 50+)					-0.17[b]	-0.16[b]
					(0.08)	(0.08)
block info. on penalty (Mincer-based)						-3.58[a]
						(1.15)
Only social housing	No	Yes	Yes	Yes	Yes	Yes
Age of babies	0	0	0-3	0-3	0-3	0-3
Observations	482	197	992	987	973	973
Pseudo R^2	0.225	0.348	0.160	0.164	0.169	0.169
Average prob.	0.43	0.50	0.50	0.50	0.51	0.51

Note: logit estimates (average marginal effects). Standard errors in parentheses with [a], [b] and [c] respectively denoting significance at the 1%, 5% and 10%. All regressions include dummies for parental education level, occupation group, *département* of residence, and years.

Table 10: The choice of an Arabic name - robustness 1

	(1) arabic	(2) arabic	(3) arabic	(4) arabic	(5) arabic	(6) arabic
one parent same name type	0.36[a]	0.35[a]	0.36[a]	0.36[a]	0.41[a]	0.34[a]
	(0.04)	(0.04)	(0.04)	(0.04)	(0.06)	(0.05)
% name type in block (aged 4-10)	0.03	0.06	0.04	0.04	0.15[c]	-0.02
	(0.05)	(0.06)	(0.05)	(0.05)	(0.08)	(0.08)
block info. on penalty	-2.95[a]	-2.23[b]	1.32	0.13	-2.29	-3.98[a]
	(1.07)	(1.13)	(1.10)	(0.36)	(1.64)	(1.54)
2nd/3rd generation, social housing	yes	yes	yes	yes	yes	yes
Specifications	Bench.	Weighted	Placebo CSP	Placebo Area	Baby girls	Baby boys
Observations	992	618374	992	986	464	470
Pseudo R^2	0.160	0.178	0.155	0.153	0.175	0.222
Mean pred. prob	0.50	0.51	0.50	0.51	0.44	0.57

Standard errors in parentheses
[c] p¡0.1, [b] p¡0.05, [a] p¡0.01

Table 11: The choice of an Arabic name - robustness 2

	(1) arabic	(2) arabic	(3) arabic	(4) arabic	(5) arabic	(6) arabic
One parent same name type	0.37[a]	0.22[c]	0.38[a]	0.33[a]	0.33[a]	0.39[a]
	(0.04)	(0.13)	(0.05)	(0.06)	(0.06)	(0.05)
% name type in block (aged 4-10)	0.06	0.17	0.09	-0.03	0.07	0.03
	(0.06)	(0.17)	(0.07)	(0.09)	(0.08)	(0.08)
block info. on penalty	-2.55[b]	-3.72	-1.47	-3.69[b]	-1.32	-7.00[a]
	(1.24)	(2.66)	(1.47)	(1.62)	(1.39)	(1.88)
2nd/3rd generation, social housing	yes	yes	yes	yes	yes	yes
Specifications	Non Mixed	Mixed	2nd gen.	3rd gen.	Non ZUS	ZUS
Observations	782	143	517	432	482	466
Pseudo R^2	0.169	0.227	0.220	0.173	0.169	0.203
Mean pred. prob	0.52	0.50	0.49	0.52	0.52	0.50

Standard errors in parentheses
[c] p¡0.1, [b] p¡0.05, [a] p¡0.01

5 Quantification of the Channels

Up to this point, we have mostly analyzed the signs and statistical significance of coefficients. Some quantification of the effect of the vertical, horizontal and economic channels in the naming decision is now provided. We first analyze the short-run effects. Then we quantify the long-run effects taking into account the dynamics of intergenerational cultural transmission.

5.1 Short run effects

In Tables 8 and 9, column (3) is our preferred specification, based on the public housing sector, and is used for quantification. All coefficients are reported as average marginal effects over choices in our sample, and are therefore easy to interpret. From column (3) of Table 8, a ten percentage point increase in the estimated occupation unemployment penalty of Arabic names translates into a 8.6 percentage points drop in the propensity to give an Arabic name. The last line of the table gives the mean predicted unemployment probability in the sample, around 19%, suggesting a quite large effect of labor market penalty. The effect is even larger for the vertical transmission motive, since in couples where at least one parent bears an arabic name, the propensity to give the same name type jumps by 23 percentage points, when a 10% increase in the share of arabic names in births of the immediate neighborhood increases probability by "only" 0.9 points. In Table 9 we see that the economic cost channel matters even more when we focus on the pure cultural transmission mechanism of migrants. Parents from a specific culture, especially a culture that conflicts with the dominant one, do matter a lot about the economic penalty inflicted to their children channel in their cultural transmission decision. A 10 percentage point increase in the block unemployment penalty reduces the probability for Arabic parents to give their baby an Arabic name by 29.5 percentage points. Conversely, the fact to have at least one of the parent with an Arabic name increases the probability of transmitting an Arabic name by 36 percentage points.

A way to quantify those effects relative to each other is to look at the model's predicted numbers of babies born with an arabic name when we shut down each of the three channels in turn. In order to calculate such counterfactuals we need to resort to the Logit structure of our econometric model where the RTA probability cannot go outside the 0-1 range. We adopt the following strategy. We start by running our benchmark Logit regression (Column (3) of Table 8), to estimate the coefficients of interest which gives us the benchmark probability of transmitting an Arabic first name in the sample. Then we select a group of observations and we run a counterfactual by attributing them other values for one or more explanatory variables. For instance we shut down the economic cost channel by forcing the *block information on penalty* to be zero for all the sample. Using the logit formula 6, we recalculate their Naming probability. This procedure ensures that the probability remains in the admissible range, while doing a "what if" experiment.

This is done in Table 12, where different lines present different scenarios. The first line reports the true number of arabic names births (700, representing more than 449'171 babies nationally). The second line is the benchmark. We then remove the vertical channel associated to parental

name. We are only able to predict 232 arabic naming decisions in that case, that is roughly a third of true births. The horizontal channel removes 84 Arabic naming decisions. The economic channel has a much stronger effect than the horizontal one: removing the economic penalty increase the number of babies receiving an Arabic name by 62%. The last line shows the results from a slightly different experiment. We make as if all blocks in the country had the same neighborhood composition and the same perceived block employment penalty. That amounts to consider the predicted number of babies when averaging the horizontal and penalty variables, which induces a slight reduction in naming choices.

Table 12: Quantification of the 3 channels

Scenario:	# babies with arabic name	weighted # babies with arabic name
true count	700	449,171
benchmark	700	447,755
no vertical (parental origin)	498	320,102
no vertical (parental name)	232	147,767
no horizontal	616	396,044
no penalty	1126	719,053
no ghetto	673	432,988

Note: This table uses logit estimates (col 3 of Table 8). Each line presents a scenario, removing in turn one of the channels of influence in the regression. Second column shows predicted number of babies born with Arabic name, in the sample of 3829 children 0-3 years old in social housing (representing 2,425,238 nationally).

5.2 Long run effects

5.2.1 The dynamic version of the model

In this section we assess the long run implications of our structural model on the naming patterns. We run various counterfactual experiments based on our structural estimates of the naming decision of descendants of migrants from Maghreb. [14] To this purpose we consider a simple dynamic extension of our structural model described in Section 3. Let consider that blocks are populated by a large number of agents $\mathcal{N}$. To keep the model tractable, we consider that agents differ only in their name type. We denote m_t the share of Arabic name holders at date t. Time is discrete. Abstracting from demographic and fertility issues, we impose a constant $\mathcal{N}$ by assuming that just before death each agent gives birth to a unique child whose name is chosen by his parent. Mortality is ruled by a Poisson process with parameter θ. The naming decision follows the model described in Section 3. We denote $(\mathbb{P}_{0,t}, \mathbb{P}_{1,t})$ the probability of giving an Arabic name for, respectively, a non-Arabic parent and an Arabic parent. Those probabilities potentially differ because of the

[14] Simulations for the population from all origins are available upon request from the authors

vertical transmission channel. The law of motion of the share of Arabic name holders is given by

$$m_{t+1} = (1-\theta) \times m_t + \theta[(1-m_t) \times \mathbb{P}_{0,t} + m_t \times \mathbb{P}_{1,t}], \tag{7}$$

Labeling μ the steady state value of m_t, we have

$$\mu = (1-\mu) \times \mathbb{P}_0 + \mu \times \mathbb{P}_1, \tag{8}$$

where the steady state probabilities of transmission, $(\mathbb{P}_0, \mathbb{P}_1)$, are characterized by equation (6). Those can be conveniently rewritten as

$$\mathbb{P}_A = [1 + \tanh(\Delta V_A / 2\sigma)] / 2, \tag{9}$$

where $\tanh(x) \equiv (e^x - e^{-x})/(e^x + e^{-x})$ and ΔV_A with $A \in \{0,1\}$ is the observable utility differential retrieved from equation (2). In ΔV_A, whatever the type of expectations (i.e. rational or backward-looking), the steady state value of the horizontal component is equal to $\mathbb{E}(m_t) = \mu$ and the parameters $(\hat{\alpha}_0, \hat{\alpha}_1, \hat{\alpha}_2, \hat{\alpha}_3)$ correspond to the point estimates retrieved from our empirical analysis where they are identified under the standard normalization assumption $\sigma = 1$. We thus have

$$\Delta V_A = \hat{\alpha}_0 + \hat{\alpha}_1 A + \hat{\alpha}_2 \mu + \hat{\alpha}_3 \mathcal{C} \text{ with } A \in \{0,1\}, \tag{10}$$

where $\mathcal{C}$ corresponds to the employment penalty of Arabic name holders.

Combining (8), (9) and (10), we obtain μ as a solution to the following fixed point equation

$$\mu = \frac{1}{2} + \frac{1-\mu}{2} \times \tanh\left(\frac{\hat{\alpha}_0 + \hat{\alpha}_2 \mu + \hat{\alpha}_3 \mathcal{C}}{2}\right) + \frac{\mu}{2} \times \tanh\left(\frac{\hat{\alpha}_0 + \hat{\alpha}_1 + \hat{\alpha}_2 \mu + \hat{\alpha}_3 \mathcal{C}}{2}\right), \tag{11}$$

First it is worth to notice that this equation does not depend on the value of the Poisson parameter θ. This makes us confident on the innocuity of our dynamic, albeit simple, demographic structure as long as we focus our analysis on the steady-state only, abstracting from any consideration on the transition dynamics. Secondly, while existence of μ follows directly from the Brouwer fixed-point theorem, uniqueness is not guaranteed and the previous equation may have multiple solutions. Contrary to Brock and Durlauf (2001, proposition 2) our dynamic setting with a non-homogenous population of agents forbids us to simply characterize the presence of multiplicity as a function of the parameter values.[15] We consequently rely on numerical computations of (11) to characterize the set of solutions $\mu(\mathcal{C})$.

[15] In absence of vertical transmission, i.e. $\hat{\alpha}_1 = 0$, our model would be included in the class of model analyzed in Brock and Durlauf (2001). Indeed, in that case, the population of babies has homogenous characteristics with respect to the naming process and our equation (11) is equivalent to their main equation (12).

5.2.2 Long run steady states and counterfactual experiments

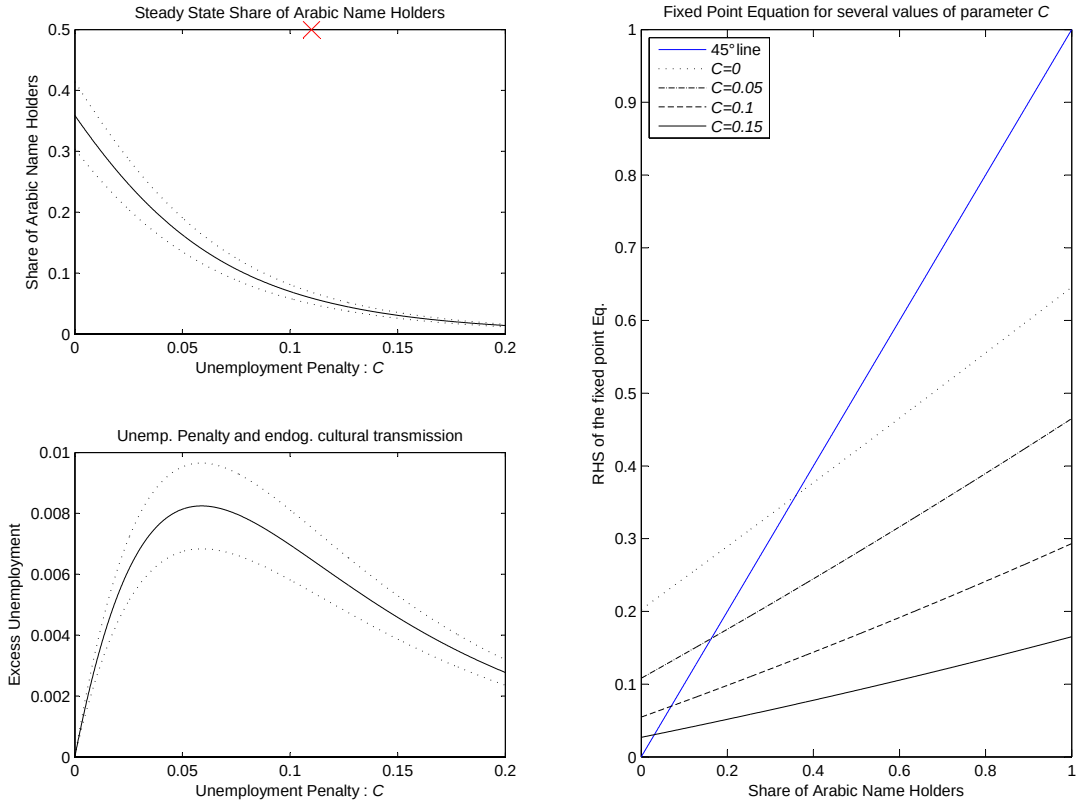
The parameters $(\hat{\alpha}_0, \hat{\alpha}_1, \hat{\alpha}_2, \hat{\alpha}_3)$ correspond to the point estimates of a specification similar to our benchmark one where the sample is restricted to descendants from Maghreb migrants in public housing (Column 3, Table 9) except that we remove all the covariates and fixed-effects which are not directly related to our structural model (5). We solve numerically the fixed-point equation (11) for values of $\mathcal{C}$ spanning the range $[0, 0.2]$; for each value of $\mathcal{C}$ this gives us the steady-state share of Arabic name holders $\mu(\mathcal{C})$. We then compute our second variable of interest, $\mathcal{U}$, which corresponds to the steady-state value of excess-unemployment due to discrimination toward Arabic name holders

$$\mathcal{U}(\mathcal{C}) \equiv \mu(\mathcal{C}) \times \mathcal{C}, \tag{12}$$

The results are reported on figure 1 for $(\hat{\alpha}_0 = -1.37, \hat{\alpha}_1 = 1.8, \hat{\alpha}_2 = 0.17, \hat{\alpha}_3 = 14.8)$. The right panel depicts the fixed-point equation (11) for various values of $\mathcal{C}$ in the range $[0, 0.2]$. We can check visually that the equilibrium is unique confirming that the horizontal channel is not large enough for generating multiple social equilibria. The top left panel reports the steady-state value of $\mu(\mathcal{C})$; the red cross represents the actual values of the unemployment penalty (equal to 0.11) and the actual share of Arabic name holders (equal to 0.5) observed in our sample of 1rts and 2nd generations of migrants from Maghreb living in public housing over the 2003-2007 period. We observe that for $\mathcal{C} = 0.11$ the steady-state share of Arabic name-holders predicted by our structural model is $\mu = 0.086$ which is smaller than the actual one. This feature might be explained by the fact that the Maghrebian migration wave is a pretty recent phenomenon in France and that most babies born in the 2003-2007 period belong to the third generation of migrants only. Hence the actual share is still far from its steady-state value and transitory dynamics are expected to bring it down in the future. We also see that in absence of discrimination ($\mathcal{C} = 0$), this steady-state share would be much larger and equal to 0.36. This confirms that the economic cost channel is also a key driver of cultural transmission in the long run.

On the bottom left panel we report the predicted excess-unemployment $\mathcal{U}(\mathcal{C})$. We observe a non-linear relationship and this confirms how a change in the degree of penalty (potentially resulting from public policy) may be partially counteracted by endogenous naming choices. Indeed, when the perceived penalty intensity ($\mathcal{C}$) falls, parents tend to raise their propensity to give Arabic names, everything else equal. This counteracting effect results in an ambiguous effect on the overall level of discrimination in the economy (total number of unemployed Arabic name holders because of the estimated penalty). Our simulation shows the interesting result that the overall discrimination starts by rising when the estimated penalty (intensity of discrimination) decreases from a high initial level.

Figure 1: Quantification of the Long-Run Effects - Descendants from Maghrebi Migrants



6 Conclusion

This paper documents the effect of economic incentives on cultural transmission in the context of naming decisions. We estimate a structural model disentangling three channels for the transmission of first names: a vertical channel from parental culture, an horizontal channel from the neighborhood culture, and an economic channel associated to the labor market penalty from holding a first name at odds with the dominant culture. We estimate those channels on the French Labor Force Survey, taking advantage of the randomized allocation of parents in public housing to analyze the spatial interactions in naming decisions within blocks. We show that the economic incentives have a first order effect in the parental decision to transmit a first name that belongs to their own culture. Parents from a culturally discriminated minority would accept an expected one year income loss for their offsprings by transmitting their cultural trait. More generally, this paper takes a first step in providing a framework to analyze the effect of public policy on the transmission of cultural values. Based on our structural model, we estimate the interplay between cultural transmission and anti-discriminatory policies on the labor and housing markets. We show that those policies might have non-trivial effects when we take into consideration the endogeneous reaction of individuals in their transmission of cultural values.

References

Adida, C., Laitin, D. and Valfort, M.A (2010): "Identifying Barriers to Muslim Integration in France" Proceedings of the National Academy of Sciences of the United States of America 107(52), 384-390.

Akerlof, G.A. and R.E. Kranton (2000): "Economics and Identity," Quarterly Journal of Economics 115, 715-753.

Algan, Y., Dustman, C., Glitz, A. and Manning, A. (2010): "The Economic Situation of First and Second Generation Immigrants in France, Germany and UK", Economic Journal 120, 4-30.

Bertrand, M. and Mullainathan, S. (2004): 'Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment Evidence on Labor Market Discrimination," American Economic Review , 94, 991-1013.

Bisin, A. and T. Verdier (2001): "The Economics of Cultural Transmission and the Dynamics of Preferences," Journal of Economic Theory, 97, 298–319.

Bisin A., G. Topa and T. Verdier (2004): "Religious Intermarriage and Socialization in the United States," Journal of Political Economy, 112, 3, 615–664.

Bisin A., E. Patacchini, T. Verdier and Y. Zenou (2010): "Bend it like Beckham: Ethnic identity and Integration", Working paper.

Blume, L. Brock, W., Durlauf, S. and Ioannides, Y. (2010): Identification of Social Interactions. in Benhabib, Bisin and Jackson, *Handbook of Social Economics*.

Breuil-Genier, P. (2001): Les variations des revenus individuels entre périodes de chömage et d'emploi, *Economie et Statistiques*, pp. 61–79.

Duguet, E., N. Landri, Y. L'Horty and P. Petit (2010): Are young French jobseekers of ethnic immigrant origin discriminated against? A controlled experiment in the Paris area *Annales dÉconomie et de Statistique*, pp. 99–100.

Ellison and Glaeser (1997)

Fernandez, R. (2008): Culture and Economics *New Palgrave Dictionary of Economics*

Fryer, R. G. and S. D. Levitt. (2004): The Causes and Consequences of Distinctively Black Names, *The Quarterly Journal of Economics*, Vol. 119(3), pp. 767-805

Glaeser and Scheinkman (2000)

Goldin, C. and M. Shim (2004): Making a Name: Women's Surnames at Marriage and Beyond, *Journal of Economic Perspectives*, Vol. 18(2), pp. 143-160

Guiso, L, P. Sapienza, and L. Zingales, (2008): Social Capital as Good Culture, *Journal of the European Economic Association Papers and Proceedings*, Vol. 6: 295–320.

Head, K. and T. Mayer (2008) "Detection of local interactions from the spatial pattern of names in France", Journal of Regional Science 48(1): 67-95.

INSEE, (2012), Fiches thématiques - Population immigrée - Immigrés - Insee Références - édition 2012 [archive], Insee 2012

Jouniaux, L. (2001): *Les 20,000 plus beaux prnoms du monde.* Hachette edts.

Lieberson, S. (2000): *A Matter of Taste: How Names, Fashions and Culture Change.* Yale University Press, 2000.

Lieberson, S. and K. S. Mikelson (1995): Distinctive African American Names: an Experimental, Historical, and Linguistic Analysis of Innovation, *American Sociological Review*, Vol. 60, pp. 928–946

Manski, C. (1993): Identification of Endogeneous Social Effects: The Reflection Problem, *Review of Economic Studies*, Vol. 60, pp. 531–542

Simon, 2003.

Tabellini, G., (2008): The Scope of Cooperation: Norms and Incentives, Quarterly Journal of Economics, 905–950.

Appendix

A Public Housing

A.1 Formal Allocation Process in Public Housings

Due to a strong "Republican ideal", the French public housing system allocates state planned moderate cost rental apartments (HLMs - *Habitations Loyer Modéré*) to natives and immigrants without concern for their cultural and ethnic background, mixing people indiscriminately. The process of allocation across public housing blocks is mainly inspired by theories from the famous architect Le Corbusier (1887-1965). Le Corbusier insisted that France must avoid the homogeneous ghettoes of the urban landscapes elsewhere, and should therefore allocate housing blind to ethnicity, not permitting family networks to grow within housing establishments. These ideas were translated into state regulation (Bernardot, 2008). This Appendix documents the legal framework for the residential allocation in the public housing sector. We show that the exogeneity of the allocation with respect to salient characteristics, such as ethnicity, is built into the law.

We first describe the eligibility criteria and the formal selection process. The only eligibility requirements for admittance into the public housing sector are to be legally living in France (as a French citizen or migrant with a valid residence permit) and to be living under a certain threshold of income per unit of consumption. This income ceiling is usually rather high: in 2009, this threshold was between 36,748 and 50,999 Euros per year for a four-person family, depending on the region of residence. As a consequence, the population eligible for public housing is on average three to four times as large as the available space in vacant dwellings. But the situation is even much tighter in the most crowded metropolis. Take the example of Paris. According to the *Observatoire du Logement et de l'Habitat de Paris* (2011), as of January 2010, there were 186,017 public housing dwellings in Paris. Public housing buildings are scattered across all Parisian areas, with a high concentration (69 percent) in six districts (the 13th, 14th, 15th, 18th, 19th and 20th *arrondissements*). Within Paris, 48.7 percent of households are under the income ceiling and could be theoretically eligible. In practice, only households with very modest incomes apply (71 percent have an income lower than the minimum ceiling for all France, equivalent to 2,345 euros per month for a household with two children). On the 31st of December 2010, there were 121,937 ongoing applications, to be compared to 12,500 public housing units allocated over the year 2010. Due to those stringent housing supply constraints, other eligibility criteria are taken into account. [16] In addition to household income, the family situation and household size are taken into account to ensure a suitable match with the characteristics of vacant dwellings, as well as the emergency of the application. Those latter criteria

[16] To apply for an apartment in a public housing sector, one has to submit a form showing the identity, the family situation, the employment status and the resources of the household, the reasons for applying to the public housing sector (currently or soon to be homeless, or reasons related to health situation, family situation, job situation, inappropriate current housing, unpleasant environment), the type of housing looked for, whether the applicant is disabled and whether this is the first application. It is important to stress the fact that the application form contains very limited information about the ethnicity of the applicant: he or she only needs to inform about his or her nationality, which is limited to three possible categories (French, European Union, or non European Union).

have recently become the main criteria used by the commission due to the boom in housing prices in the private sector during the mid-90s and the 2000s. In particular, five priority criteria are defined by law (Article L441-1 of law relative to construction and housing - *Code pour la Construction et l'Habitat*) at the national level to ensure that vacant housing will first be distributed to households with obvious social difficulties. Households satisfying these priority criteria are those in which there is a (mentally or physically) disabled person, those living in precarious or hazardous shelter due to financial constraints, those living in a temporary accommodation, individuals living in a precarious shelter who recently found a job after a long unemployment spell, and spouse-abused individuals.

Regarding the selection process, the commissions of selection in charge of allocating households to vacant public housing dwellings are held at the *département* level (or at the city level in the case of Paris which is both a city and a *département* due to its size).[17] The composition of the commissions is regulated by law: it includes six members of the public housing offices board, a representative of associations for social and economic insertion (appointed by the head of the *département*), mayors of the cities (or districts) in which vacant housings are to be attributed, as well as a representative of any association defending tenant rights. In addition, another *département* representative may attend the commission. For each vacant housing unit, at least three households must be considered by the commissioners, who finally decide which household will be allocated to which housing unit, according to the eligibility and priority criteria detailed above. Other criteria such as the number of children in the household are also taken into account in order to allocate suitable dwellings.

Despite this legal process of allocation, one might still be worried about the possibility of self-sorting of households that refuse the residential allocation proposed by the commission. In theory, households can refuse up to three offers. However, self-sorting, especially on ethnic characteristics, seems unlikely to be a common practice. Residential mobility within the public housing sector is very low, due to the current strong shortage of supply of public housing dwellings. This makes it unlikely that the selected households could be really picky about the diversity of their neighborhood (see the study by Simon, 2003). In addition, rents are considerably lower in public housing than in private housing, increasing the opportunity cost of moving, so that the turnover is very low. More specifically, the mobility rate in the public housing sector is even lower than for recent owners. Public housing allocation in Paris serves as a useful concrete example. The mobility rate (defined as the ratio of new entrants over the total number of public housing dwellings) is particularly low: it reaches 5.5 percent in 2010. It is formally possible to indicate a precise neighborhood in the application form, but in practice, very few applicants (6.6 percent) do provide this information. More than half of the 121,937 applicants (52.9 percent) did not mention any particular area at all, probably due to the fear of being rejected on this ground. Among those who indicated an area of preference, 91.2 percent mentioned the area where they were already living. People who move

[17] Metropolitan France is divided into 22 large administrative areas, called *régions*, and into 96 smaller administrative areas, called *départements*. Each *département* is hence a subdivision of a region, and several *départements* can belong to the same region. Each *département* is administered by an elected General Council (*Conseil Général*) and its President, whose main areas of responsibility include the management of a number of social and welfare programs, junior high schools, buildings and technical staff, local roads, schools, rural buses, and municipal infrastructure.

within the public housing sector are people who moved for larger space following an increase in their household size (only 12 percent of the public housing dwellings have more than three rooms).

A.2 Descriptive statistics on public housing

Table 1 shows the summary statistics of the public housing units in the Labor Force Survey. The survey comprises 1,535 public housings, with an average of 45.12 public housing blocks by *département*. The average number of households within a block is 17.99, which is similar to the whole sample. We observe 618 new born babies in the social housing sector, which will be used as our sample for the estimates centered on social housing.

Table 13 reports the descriptive statistics of the sample of individuals aged over 15 years and who are consequently interviewed in the LFS. There is an over-representation of Arabic Name holders in public housings: the share reaches 14 percent in this type of housing against 4 percent in the total sample. Similarly, the proportion of individuals who have a Maghrebian nationality at birth, or who have at least one parent with a Maghrebian nationality at birth, is 16.8 percent in social housings. This proportion is three times as high as in the total sample (5.61 percent). Individuals who enter the public housing sector have also lower socio-economic backgrounds than the rest of the population. The share of unemployed is almost twice as high in public housings (9.6 percent) than in the total sample (5 percent). There is an over-representation of blue-collars and clerks (78 percent in the social housing, and 55 percent in the whole sample) and an under-representation of executives (3 percent in social housing, 12 percent in the whole sample). This table shows that there is a clear selection of individuals into the public housings, since the eligibility is based on the socio-economic characteristics. However, our key identification strategy relies on the exogeneous allocation of individuals.

A.3 Additional tests on the Exogeneous Spatial Allocation Process in Public Housing

A.3.1 Absence of self-sorting on ethnic backgrounds

Our first set of alternative tests consists in showing that while households tend to self-segregate in the unconstrained private housing market, there is no such evidence in the public housing market. We test this using the LFS and focusing on individuals who recently moved into an area (within the previous year).

We first estimate the correlation between the origin (nationality) of individuals moving into a new area and the share of the area's "long term" population of the same origin.[18] We expect

[18] A similar test was proposed by Goux and Maurin (2007) to show that the educational achievement of the children of newcomers in public housing is uncorrelated with that of the current residents. Individuals do not self-select in public housing neighborhoods according to the educational achievement of the neighbors' children. By contrast, the authors find a strong self-selection on the educational characteristics in the private housing sector.

Table 13: Descriptive statistics of individuals

	Total sample	Public housing
	Mean (std)	Mean (std)
Age	45.22 (19.22)	40.93 (18.19)
Gender (Male)	0.48 (0.49)	0.46 (0.49)
Married	0.61 (0.48)	0.49 (0.50)
Arabic names	0.04 (0.21)	0.14 (0.35)
Maghrebian Nationality at Birth	0.05 (0.23)	0.16 (0.37)
Employed	0.52 (0.49)	0.50 (0.49)
Unemployed	0.05 (0.20)	0.09 (0.27)
Inactive	0.43 (0.49)	0.41 (0.49)
Hourly wage (euros)	9.70 (4.29)	8.24 (3.00)
Occupation: executive	0.12 (0.33)	0.03 (0.18)
Occupation: intermediate	0.20 (0.40)	0.14 (0.34)
Occupation: clerk	0.30 (0.46)	0.40 (0.49)
Occupation: blue collar (skilled)	0.15 (0.36)	0.20 (0.40)
Occupation: blue collar (unskilled)	0.10 (0.30)	0.18 (0.39)
Occupation: craftman	0.06 (0.24)	0.02 (0.14)
No education	0.21 (0.41)	0.36 (0.48)
Elementary school	0.44 (0.49)	0.42 (0.49)
High school	0.14 (0.34)	0.11 (0.31)
College	0.09 (0.28)	0.05 (0.22)
Graduate	0.10 (0.30)	0.04 (0.20)

a significant relationship in the private housing market where location choice is relatively unconstrained but not in the public housing sector. Table **??** reports the results from an OLS regression of the share of neighbors from the same origin as new movers on new movers' characteristics: nationality group, public housing dummy, quadratic function of age, hourly wage (log) education, socio-economic category, *département* fixed effects, and interaction of individual characteristics with the public housing dummy. We consider seven different nationality groups: native French, naturalized French, Europeans, Maghrebians, other Africans, Asians, and other nationalities, which is taken as the reference group.

Three facts are worth noting here. First, there is indeed evidence that on average native French are significantly more likely to move in neighborhoods where the share of natives is higher, compared to households from other nationalities. This is not surprising given the fact that natives make up a large majority in the French population. The second interesting point is that the coefficient for living in the public housing sector is negative and significant at the 5% level. More precisely, it reveals that HLM households move in areas where the share of individuals from the same origin is on average 15.5% lower than for households in the private housing sector. This result strengthens the idea that the extent to which households in the public housing sector live close to their co-ethnics is lower than in the private sector. Finally, when we interact nationalities with the public housing dummy, none of the coefficients are statistically significant. This comforts us with the idea that there is no particular self-segregation along ethnic lines in the public housing sector. We have run the same kind of test on other individual characteristics, and reach similar conclusions. We find that in the private sector, highly educated or low skilled workers are very likely to move into neighborhoods with higher levels of highly educated (respectively low skilled) people. This is not surprising and illustrates self segregation along education level in the private sector. On the contrary, such segregation does not appear in the public housing sector.

A.3.2 Tests on the refusal rate of public housing offers

The previous tests point out the absence of self-selection along ethnic lines among the movers. But self-selection could occur prior to the move. In this case the sample of movers that we observe in the database would be biased. We address this issue by looking at households that have refused a public housing dwelling offer. We show that even if households declined at least one offer, possibly due to the ethnic diversity or the socio-economic composition of the neighborhood, they were still unable to choose the ethnic and socio-economic composition of the housing block in which they end up living. In this section, we show the results for the refusal rate of public housing offers depending on the ethnic composition of the neighborhood. But we find similar results when looking at the composition of the housing block by education and occupation characteristics.

We run this analysis on an alternative database, the 2002 Housing Survey, to bring additional evidence on the absence of sorting. The Housing Survey (HS) shares exactly the same structure as the Labor Force survey, with information collected at the housing block level with adjacent neighbors (an average of 18.2 in the public housing sector). But the HS also reports specific

questions on household satisfaction with housing quality and if they have turned down public housing offers.

First, if there were self-selection upon diversity, we should expect households that turned down proposals before being allocated to their current public housing dwelling to end up living in less diverse neighborhoods. To test this conjecture, we run OLS regressions of a variable indicating whether the household declined at least one offer (during the latest application process) on the level of diversity of the neighborhood in which it now lives. Panel A-I of Table 15 shows various estimates of the effect of ethnic diversity on the probability of having turned down offers. Column 1 shows the correlation without any additional control variables. In Column 2, we control for household characteristics. We add up the characteristics of the housing project in Column 3. Column 4 finally includes neighborhood characteristics and *département* fixed effects since the allocation of a public housing dwelling takes place at the *département* level. In each specification, the coefficient on ELF is not significantly different from zero, showing that households having declined offers during their past allocation process do not end up living in neighborhoods with significantly different levels of diversity.

We explore further the validity of this conjecture by focusing on the reasons adumbrated by households for refusing an offer. If public housing residents were to sort themselves on the basis of their (dis)taste for diversity, those who declined "because of the local environment" should now live in significantly less diverse neighborhoods. We thus regress a dummy variable indicating whether an "unpleasant environment" was the reason why the household declined at least one offer (during the past application process) on the level of diversity of its current neighborhood. Panel A-II of Table 15 reports the estimates on the level of diversity, using the same specifications as above. Here again, none of the coefficients is significant. Instead, household characteristics such as the labor market status of the head of household and the size of the household are the only ones that matter in these regressions.

Alternatively, we perform these tests on the subsample of individuals currently waiting for a public housing offer. Panel B of Table 15 shows the regressions of the refusal dummy (B-I) and the "refusal due to unpleasant environment" dummy (B-II), for individuals who are currently applying for public housing on the diversity in their neighborhood. We still control for household, building and neighborhood characteristics. Ethnic diversity of a block is uncorrelated with households wait-listed for a public housing assignment having turned down offers since the beginning of their request (B-I). This suggests that the current level of diversity in the block does not rush households out of the area, as their propensity to decline an offer is independent of the ELF in the current neighborhood. The high refusal rates of public housing offers do not therefore seem driven by a hope to reduce diversity by waiting.

Let us now focus on individuals who left their previous housing unit because they did not like the environment. In the Housing Survey, 5 percent of households that moved over the past four years mention an unpleasant environment as one of the main reason they moved. In this question, the

phrase "unpleasant environment" explicitly refers to troubles such as "noise, lifestyle or insecurity". Again, this could be related to high levels of diversity. If this is true, and if households can actually select the block to which they move, then we expect that those households having moved because they disliked their environment ended up living in less diverse neighborhoods than the households that moved for a different reason.

We perform OLS regressions of a variable indicating whether the household left its previous housing due to an unpleasant environment, on the level of diversity of its current neighborhood. Table 16 shows the coefficients on diversity in the specification controlling for household, building and neighborhood characteristics, and including *département* fixed effects. Column 1 shows the results for households that moved within the private housing market. As expected, households that left their previous housing to escape from an unpleasant environment now live in blocks where the diversity is significantly lower. Column 2 shows that this result does not hold for households that moved within the public housing market. This result suggests once again that in the public housing sector, households do not have control over the diversity of the block to which they are allocated.

A potential concern with the previous result is due to the small sample of observations (only 627 in the public housing case), generating large standard errors. Therefore, we replicate this test on a larger subsample. Instead of focusing on households that have moved within a housing sector, we now concentrate on households having moved into each sector, no matter the sector in which they were living prior to their move.[19] As previously, we see that for households living in the private housing sector, the probability that they left their previous housing due to an unpleasant environment is negatively correlated with the diversity in the current neighborhood (Column 3). However, no such significant relationship shows up for households living in a public housing dwelling (Column 4), and the estimates are now more precise than in Column 2. We can infer from those tests that households tend to self-select in low-diversity neighborhoods in the private housing sector, but are unable to do so in the public housing sector.

[19]To summarize, Columns 1 and 2 report the results for households moving from a housing dwelling in the private (1) and public (2) sectors into a housing dwelling in the same sectors. Columns 3 and 4 report the results for households moving from any housing sector into the private (3) and public (4) sectors.

Table 14: Correlation between new inhabitants' nationality and share of the housing block population of the same nationality

Dep Var: Share of population of the same ethnic group as new movers in a given housing block	
Nationality (reference group: Other nationalities)	
Native	0.067**
	(0.030)
Naturalized French	-0.037**
	(0.012)
European	-0.007
	(0.011)
Maghrebian	0.007
	(0.013)
African	-0.016
	(0.015)
Asian	-0.019
	(0.047)
Public Housing (HLM)	-0.184**
	(0.066)
Nationality * HLM	
HLM * Native	0.040
	(0.033)
HLM * Naturalized	0.051
	(0.036)
HLM * European	0.010
	(0.037)
HLM * Maghrebi	0.024
	(0.036)
HLM * African	0.017
	(0.039)
HLM * Asian	0.027
	(0.078)
Intercept	0.070**
	(0.035)
R-squared	0.864
N	11519

Note: Additional controls are a quadratic function of age, gender, hourly wage (in log), education, occupation, housing block socio-economic characteristics and department fixed effects.
* p¡0.10, ** p¡0.05, *** p¡0.001

Table 15: Rejection of HLM offers and Ethnic diversity

Rows: Dependent Variables	Coefficient associated with Ethnic Diversity (1)	(2)	(3)	(4)
Panel A: Sample of households who currently live in public housing:				
I. Probability of having declined at least one HLM offer during the previous application process	0.058 (0.058)	0.069 (0.063)	0.017 (0.067)	0.123 (0.0886)
N	1,779	1,779	1,748	1,744
R^2	0.001	0.021	0.023	0.089
II. Probability that the reason for having declined an HLM offer during the previous application was "unpleasant environment"	0.162 (0.144)	0.061 (0.158)	0.017 (0.171)	-0.0310 (0.258)
N	417	417	415	414
R^2	0.003	0.035	0.050	0.308
Panel B: Sample of households who are currently applying to public housing:				
I. Probability of having declined at least one HLM offer during the current application process	-0.063 (0.057)	-0.043 (0.064)	-0.088 (0.071)	-0.116 (0.103)
N	1,192	1,192	1,173	1,171
R^2	0.001	0.014	0.024	0.121
II. Probability that the reason for having declined an HLM offer during the current application was "unpleasant environment"	0.004 (0.194)	-0.007 (0.237)	-0.104 (0.250)	-0.122 (0.506)
N	198	198	195	194
R^2	0.000	0.083	0.115	0.590

Each of the coefficients is estimated from a separate regression of each of the four dependent variables described in the first column on ethnic diversity. Column 1 does not include any control. Column 2 includes households characteristics (gender, age, education, employment status and nationality of the head of household, total income (in log) of the household per unit of consumption, and household size). Column 3 adds up the characteristics of the building (number of apartments (in log) and construction date). On top of that, column 4 includes neighborhood characteristics (socio-economic background (Tabard index), and local unemployment rate), as well as *département* fixed effects. In addition, a dummy variable indicating whether the household already lives in the public housing sector is included in specifications 2 to 4 of Panel B. The coefficients for all the controls are available from authors upon request. Robust standard errors adjusted for block clustering are in parentheses. *** p¡0.01, ** p¡0.05, * p¡0.1

Table 16: Do households having left their previous housing due to an "unpleasant environment" now live in less diverse neighborhoods?

Dependent Variable:	Main reason for leaving previous housing: unpleasant environment (noise, lifestyle or insecurity)			
	Households who moved **within** the		Households who moved **toward** the	
	Private Housing sector	Public Housing sector	Private Housing sector	Public Housing sector
	(1)	(2)	(3)	(4)
Ethnic Diversity	-0.073**	0.083	-0.061*	0.016
	(0.030)	(0.140)	(0.032)	(0.052)
Observations	5,955	627	6,560	1,793
R-squared	0.030	0.207	0.031	0.079

In each regression, we control for household characteristics (gender, age, education, income (in log), employment status, nationality, household size), building characteristics (number of apartments and construction date), neighborhood characteristics (socio-economic background (Tabard index), and local unemployment rate) and *département* fixed effects. The coefficients for all the controls are available from authors upon request.

Robust standard errors adjusted for block clustering are in parentheses. *** p¡0.01, ** p¡0.05, * p¡0.1