THE ECONOMIC JOURNAL

The Economic Journal, **132** (*February*), 437–470 https://doi.org/10.1093/ej/ueab058
Advance Access Publication Date: 13 July 2021

THE ECONOMIC INCENTIVES OF CULTURAL TRANSMISSION: SPATIAL EVIDENCE FROM NAMING PATTERNS ACROSS FRANCE*

Yann Algan, Clément Malgouyres, Thierry Mayer and Mathias Thoenig

This paper studies how economic incentives influence cultural transmission, using a crucial expression of cultural identity: child naming decisions. Our focus is on Arabic versus non-Arabic names given in France over the 2003–2007 period. Our model of cultural transmission features three determinants: (*i*) vertical (parental) cultural transmission culture; (*ii*) horizontal (neighbourhood) influence; (*iii*) information on the economic penalty associated with Arabic names. We find that economic incentives largely influence naming choices: if the parental expectation on the economic penalty was zero, the annual number of babies born with an Arabic name would be more than 50% larger.

Cultural traits persist across generations, partly because individuals exhibit a preference for the transmission of their own culture to their offspring.[1] However, external forces might be operating through both social pressure and the economic environment that restrict how the individual desires to transmit ones' culture translates into actual choices. These constraints are likely to be especially binding for minorities. This is naturally true of immigrants who live in societies in which natives tend to value conformity/assimilation and express anxiety with respect to (actual or perceived) rising cultural diversity.[2]

In this paper, we analyse how social and economic forces constrain the inter-generational transmission of culture among immigrants and their descendants with particular emphasis on the tension between the taste for the perpetuation of inherited cultural traits and the perceived

* Corresponding author: Yann Algan, Economics Departement, Sciences-Po, 28, rue des Saints-Pères 75007 Paris, France. Email: yann.algan@sciencespo.fr

This paper was received on 21 December 2017 and accepted on 6 July 2021. The Editor was Hans-Joachim Voth.

The data and codes for this paper are available on the Journal website. They were checked for their ability to reproduce the results presented in the paper. The authors were granted an exemption to publish parts of their data because access to these data is restricted. However, the authors provided a simulated or synthetic dataset that allowed the Journal to run their codes. The synthetic/simulated data and the codes for the parts subject to exemption are also available on the Journal website. They were checked for their ability to generate all tables and figures in the paper; however, the synthetic/simulated data are not designed to reproduce the same results.

We thank Alberto Alesina, Alberto Bisin, Pascaline Dupas, David Laitin, Paola Giuliano, Laurent Gobillon, Camille Hémet, Raquel Fernandez, Rachel Kranton, Luigi Pistaferri, Nancy Qian, Yves Zenou and Fabrizio Zilibotti for their helpful comments. We also thank the participants of seminars and conferences at Stanford, Berkeley, RSAI in Miami, the NBER Summer Institute, 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, ENSAI in Rennes, Universities of Bern, Lausanne and Lugano for their helpful comments. Yann Algan acknowledges financial support from the ERC Starting Grant 240923. Clément Malgouyres thanks the support of the EUR grant ANR-17-EURE-0001. Mathias Thoenig acknowledges financial support from the ERC Starting Grant GRIEVANCES-313327.

[1] See Bisin and Verdier (2011) for a review of theoretical and empirical studies on the inter-generational transmission of culture.

[2] For instance, Hainmueller and Hiscox (2010) provided evidence suggesting that non-economic factors are important drivers of hostility to immigration. Hainmueller and Hopkins (2014) surveyed the literature and highlighted the dominance of concerns over cultural impacts in shaping public attitudes toward immigration.

economic discrimination attached to them.[3] We study the determinants of this trade-off for a specific case of transmission: the cultural type of first names parents give their children. We focus on the decision of whether to give a first name associated with Arabic/Muslim culture to babies born in France in the early 2000s. We view this decision as an appropriate object of inquiry for two main reasons: (1) first names are widely considered important markers of cultural identity—the choice of a first name is available to all parents, without material constraints, and is thus sometimes referred to as a 'pure' expression of cultural identity (Lieberson, 2000); (2) there can be direct economic consequences to naming decisions.[4] Several studies have shown that first names associated with a cultural minority are perceived negatively by employers (Bertrand and Mullainathan, 2004).

In the French context, Arabic name holders are associated with both severe economic discrimination and with cultural elements that potentially conflict with the 'traditional' (native) culture: religion, migration, political tensions, historical legacy with ex-colonies or even consumption habits. First, the largest immigration wave since 1945 originates from former North African colonies (mostly Algeria, Morocco and Tunisia, three countries that we refer to as Maghreb). Accounting for first-generation migrants and their descendants, this migration wave represents approximately 5.7% of the current French population (INSEE, 2016). Second, the decolonisation process was conflictual for those countries—particularly Algerian independence that occurred after several years of violent war (1952–1964). Third, Arabic names are also a sign of the Muslim religion since most of those names come from the Quran, and the transmission of first names associated with the Quran is a natural practice for religious people. By contrast, non-Arabic names in France are mainly associated with Saints' names, i.e., coming directly from the French calendar of Christian Saints (or inspired by it). Fourth, economic prejudice against Arabic name holders has been largely documented in France. Second-generation migrants from Maghreb face the highest penalty on the French labour market among the different immigrant groups (Algan *et al.*, 2010; Duguet *et al.*, 2010). Combes *et al.* (2016) used the Labour Force Survey to show that Arab immigrants working in customer-facing jobs have much higher risks of unemployment. Even closer to our main variable of interest, Adida *et al.* (2010) performed an audit study using CVs that only differ with respect to the origins of the first name (Arabic versus Christian). Vitas with an Arabic first name are 2.5 times less likely to receive a job interview callback compared to their Christian-named counterparts, everything else being equal, including the last name. Consistent with this recent work on French data, we document a large penalty attached to Arabic names on the French labour market: the average unconditional differential of unemployment between Arabic name holders and the rest of the population amounts to 13% in our sample.

Our empirical design is based on a random-utility discrete-choice model of parental naming decisions. The choice is binary and pertains to the cultural type of the child's first name, Arabic or non-Arabic. The model incorporates the two traditional vertical and horizontal channels

[3] We adopt here a popular definition of culture among economists, as being the belief, behaviour or identity patterns that are transmitted from one generation to another (Fernández, 2011). Recent surveys include Algan and Cahuc (2013) and Alesina and Giuliano (2015).

[4] Fryer and Levitt (2004) have provided additional evidence on the cultural component of first names by showing that the surge in distinctively Black names in the United States since the seventies could be associated to a rise in Black cultural identity. In their study of two major waves of immigration in the United States, Abramitzky *et al.* (2020) emphasised the attractiveness of first names as a measure of assimilation. They argued that first names are more likely to reflect preferences and less likely to reflect constraints imposed by the host society than alternative measures, such as inter-marriage—which could reflect both the demand and supply determinants of assimilation opportunities.

analysed in the literature on cultural transmission (Bisin and Verdier, 2001), to which an economic channel is added. The vertical transmission channel results from the utility gain for parents when transmitting their own cultural type. The horizontal transmission channel stems from spatial externalities associated to the cultural types of peers and neighbours. The economic channel corresponds to the expected economic penalty inflicted on one's children when giving an Arabic first name.

The French Labour Force Survey (LFS henceforth) provides a unique source of information for measuring and estimating these various cultural transmission channels. The vertical transmission channel is identified by contrasting the first names of parents and children, all being reported in the survey. Regarding the measurement of the economic channel, the LFS allows for the detailed computation across occupations of unemployment rates associated with Arabic/non-Arabic names. Finally, the LFS data collection is based on a large representative set of more than ten thousand sampling units spread all over the country, each unit consisting of a residential block of *twenty adjacent households, all of which are surveyed*. This feature enables us to define a set of relevant peers at the very local level where many social interactions have been shown to occur in France (see Goux and Maurin, 2007 and Maurin and Moschion, 2009). We use this set of peers for two purposes: first, to measure the horizontal transmission of naming choices from nearby neighbours; second, we compute the *local information on penalty* (LIP hereafter) as the average unemployment differential between Arabic and non-Arabic name holders across these neighbours' occupations. The idea underlying this key explanatory variable is that parents use various sources of information to form a belief about the economic penalty associated with Arabic names and our hypothesis is that one of the main sources of information is people living in the same neighbourhood. This mechanism is likely to be especially relevant for migrants (and their descendants), since they should exhibit a low initial knowledge of the local labour market (see Hellerstein *et al.*, 2011 and Goel and Lang, 2019 for relevant evidence on Canadian and US labour markets).

A critical issue in estimating our model relates to parents' endogenous location choices across residential neighbourhoods, resulting in spatial sorting on—possibly unobserved—characteristics, correlated with the propensity to give Arabic first names to their offspring. It could be, for instance, that parents most attached to transmitting an Arabic name to their children choose to live in residential blocks with religious neighbours, who themselves tend to work in low-discrimination occupations. We mitigate this concern by restricting our estimation to a sample of households living in the French public housing sector. Due to legal and binding dispositions, state-owned apartments are allocated to households without consideration for their cultural background, mixing people indiscriminately. Furthermore, individuals rarely move since the rents are much lower than market rates. Building on Algan *et al.* (2016), we confirm, with a variety of tests, that spatial allocation within the public housing market can be considered as good as random.[5]

Our main result is that economic factors are important drivers of individual cultural transmission decisions. We find that an increase in the perceived penalty associated with Arabic name

[5] We also consider the possibility that parents retrieve information from the unemployment differential they observe in their own occupation. However, while our focus on public housing in the French context allows us to deal with spatial sorting, there is no similar device to avoid parental sorting across occupations. We therefore prefer the neighbourhood-based approach that, as detailed in Section 2, features fixed effects for the parental occupation and thus accounts for the fact that parental occupation is not random.

holders, as measured by our LIP variable, reduces the probability that parents will give such names. The magnitude of the effect implied by our estimates is also quite sizeable: if the parental expectation of the economic penalty were brought down to zero, the annual number of babies born with Arabic names in France would be more than 50% larger. In terms of the two other channels, the vertical channel is by far the dominant factor in the naming decision: a French baby who has at least one parent or grandparent with an Arabic/Muslim background is twice as likely to be given an Arabic name. The horizontal channel is statistically significant in some regressions, but quantitatively much less important. While these findings hold for the sample of all households living in public housing, they are mainly driven by the behaviour of first and second generations of migrants from Arabic countries. Using our theory-grounded estimates, we are also able to quantify welfare gains and losses attached to cultural transmission. Focusing on the substitution rate between the vertical and the economic cost channels, we can express the strength of cultural attachment in monetary units. For first and second generations of migrants, we find that vertical transmission of an Arabic name provides the same shift in parents' utility as a 3% rise in the child's lifetime income. Finally, we also assess the welfare effects of French policies historically aimed at restraining naming choice.

Our paper fits into several strands of research. A substantial body of work by economists studies the transmission of cultural values and the formation of identity (Akerlof and Kranton, 2000; Shayo, 2009; Atkin *et al.*, 2019). Bisin and Verdier (2001) provided a seminal cultural transmission model, distinguishing between vertical transmission by parents and oblique or horizontal transmission associated with social interactions. Tabellini (2008) and Guiso *et al.* (2008) modelled the interactions between norms and economic incentives in the inter-generational transmission of values like trust. Bisin *et al.* (2004) and Bisin *et al.* (2016) estimated structural models of transmission of religious values and ethnic identity. We contribute to this literature by introducing a new channel of cultural transmission through economic incentives. We also innovate in our empirical application in terms of measurement, since we observe variation in incentives at the block level as opposed to more aggregate units. Most important, building on Algan *et al.* (2016), we exploit the quasi-random allocation of households across blocks among public housing tenants as an identification strategy.

The stream of recent work studying the question of migrants' assimilation and how discrimination affects it is perhaps the most directly relevant to our findings. Abramitzky *et al.* (2016) analysed the co-evolution of cultural and economic assimilation during the age of mass migration in the United States: they found significant first-name assimilation among immigrants that tends to translate into better economic outcomes for their offspring.[6] Mazumder (2019) found that immigrants' military service in the US army during World War I increased their rate of cultural assimilation, with potentially positive economic returns. Most closely to our study, Fouka (2019) found that immigrants from German origins responded to discrimination during World War I in the United States by increasing their assimilation efforts—partially by changing the 'Americanness' of their names. Our results are consistent with Fouka (2019) since they show that minority parents are willing to undertake costly assimilation actions when exposed to more information about discrimination. This could indicate that the discrimination at play is to some extent conditional, meaning that it is likely to be lower if individuals send signals of loyalty

[6] In a subsequent version of their paper, Abramitzky *et al.* (2020) compared first-name assimilation across two periods of intense immigration in the United States.

to the dominant culture (Bisin *et al.*, 2011).[7] Our contribution with respect to this literature is two-fold. First, we analyse a channel of cultural transmission directly associated with a measure of discrimination on the labour market. Second, by focusing on parents who are exogenously allocated to their neighbours within public housing, we are able to exploit fine-grained exogenous variation in neighbourhood ethnic and occupational composition to estimate jointly the vertical and horizontal channels as well as the economic cost—whereas most of the previous literature focuses on the heterogeneous effects of aggregate shocks.

Our paper also relates to the literature on the link between long-run economic and cultural change. First, we document a very high preference for vertical transmission among parents with Arabic origins, in line with the vast literature highlighting the strong persistence of cultural norms (see Guiso *et al.*, 2016 for instance). Faced with adverse economic consequences of vertical transmission, we find, however, that parents adjust their behaviour. This evidence is consistent with the notion that culture tends to be a highly persistent construct that can nevertheless evolve in the face of changing circumstances—a fact that has been well documented in the literature on long-run persistence. For instance, Nunn and Wantchekon (2011) showed that the intensity of exposure to the slave trade in Africa is associated with a lower level of trust nowadays due to an impact on cultural norms and values. Relatedly, Voigtländer and Voth (2012) studied the persistence of anti-semitism in Germany from the Middle Ages onward. They documented that the persistence of anti-semitism, while high on average, is much lower in cities with a history of demographic expansion and exposure to economic exchanges (e.g., trade).

Finally, our paper relates to the literature on discrimination and its consequences for economic decision and public policy. A fairly large body of research has found evidence of discrimination by employers against first names from a cultural minority. Those studies exploit either audit study methodology (Bertrand and Mullainathan, 2004; Adida *et al.*, 2010; Duguet *et al.*, 2010) or representative surveys (Heckman, 1998; Fryer and Levitt, 2004). Our paper differs from this literature in two main respects. First, our ambition is not to measure economic discrimination, but to analyse how parents react to perceived *information* on discrimination. We propose different information channels that parents can retrieve from the labour market and estimate whether they use this information in their cultural investment. Second, we analyse the ex ante decision of adopting (or not) a cultural trait that could be discriminated against, while the rest of the literature focuses on the ex post economic consequences of carrying this cultural trait. We show that the determinants of cultural identity, and more specifically the economic ones, have substantial welfare and public policy implications that have been overlooked so far.

The remainder of the paper is as follows. Section 1 provides a detailed description of the data we use. Section 2 presents our theoretical model of naming decision. Section 3 contains our baseline estimation results. Section 4 provides various robustness tests. Section 5 quantifies the

[7] Naturally, whether the lower transmission of identity through name giving is mostly a strategy or a true shift in beliefs is an open question. In a closely related paper, Fouka (2020) showed that the second generation of German-origin Americans facing German language bans while at school in the United States following World War I displayed signs of lower assimilation. In particular, they were more likely to marry within their ethnic group (Germans) and to give German-sounding names to their children. This suggests that some form of discriminatory public policies might backlash and feed oppositional identities (consistent with the findings of Lleras-Muney and Shertzer, 2015 and Rozenas and Zhukov, 2019). At any rate, given the documented labour market penalty associated with Arabic names, the response we identify is likely to have long-run real implications on the economic well-being of newly born children even if name assimilation does not reflect a pure convergence in beliefs.

Table 1. *Descriptive Statistics of the Residential Blocks.*

	All	Public housing
Number of blocks	10,528	2,674
Number of blocks by department	108.54	28.45
Average number of households per block	16.31	17.59
Average number of members per household	2.43	2.51
Average number of children per household	0.50	0.65
Total number of households	173,154	26,958
Total number of individuals	425,223	69,437

contributions together with welfare effects of the vertical, horizontal and economic channels on cultural transmission.

1. Data

1.1. *The French Labour Force Survey*

Our empirical analysis is based on the French LSF from 2003 to 2007. The LFS is a representative survey of the French population, stratified at levels of around 3,500 residential blocks per year, with each block defined as an average of 20 adjacent households. The LFS is a rolling panel of six quarters and all the households within a given block are interviewed every quarter. All household members older than 15 years are interviewed, and they report information on their socio-economic characteristics, including employment status (unemployed, inactive and employed), hourly wage and 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 twenty-nine 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 household members, including children below 15 years old.

Since the data collection is based on (very) close neighbours, the LFS provides a unique opportunity to understand the role of horizontal factors in the transmission of names. Given that the sampling unit in the LFS consists of small groups of adjacent households, and that all the members of the households within the same block are interviewed, we get detailed information on all individuals living in the neighbourhood. Another important characteristic of the LFS is that it distinguishes between the public and the private housing sectors. As discussed below, our identification strategy will be based on residential allocation of households within the public housing sector. Thus, we report on both the total sample and on the sample of public housing residential blocks.

The time span of the rolling panel is too short (six quarters) to exploit time variation in the socio-economic composition within residential blocks. Thus, we keep one observation per member of the household, which generally corresponds to the first wave of interviews. Table 1 reports the main descriptive statistics of the full database when we use this selection criterion. Our total sample is made up of 10,541 blocks, with 1,535 blocks containing one public housing unit. The average block size is 18.31 adjacent households, and each household consists of 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.

Table 2. *Transmission of Name Types.*

	Babies with:	
	non-Arabic name	Arabic name
Newborn (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 generations):		
Parents with non-Arabic name	461 (183)	111 (47)
Parents with Arabic name	658 (317)	789 (461)

Notes: This table reports the number of babies by name type and allocates them according to the name type of their parents. The top panel gives figures for the whole sample of babies born within the year. The bottom panel considers babies aged 0 to 3 at the time of the survey, born from at least one parent or grandparent with Arabic origins. The sample of babies in the public housing sector are given in parentheses. The term generation refers to the generation of the babies (e.g., a second generation baby is the child of a first generation migrant).

1.2. *Sample of Babies' Names*

Our main variable of interest is the individuals' name type and the cultural background that is associated with it. We focus on the transmission of Arabic first names, as opposed to non-Arabic names, in French society. In our data, we code Arabic first names according to the classification of Jouniaux (2001). Arabic names are associated with the most important population of immigrants in France—Maghreb—and to a lesser extent with the Middle East (other Arabic countries and Turkey), in the aftermath of decolonisation initiated in the 1960s. According to INSEE (2016), people with Maghreb origins (i.e., Algeria, Morocco, Tunisia) represent almost 60% of non-OECD migrants from first and second generations in France in 2008; this corresponds to 3.7 million individuals (1.7 million for the first generation and 2.0 million for the second generation) out of a total French population of 64.3 million.

We describe in Table 2 our sample of babies along three dimensions of relevance for our empirical analysis: (*i*) the cultural type of parents' first name; (*ii*) the cultural background of babies (as captured by the immigration history of the household); (*iii*) whether the household lives in private or public housing. We start in the upper panel with the full sample of 3,541 newborn babies over 2003–2007 for whom we have all the needed information on the parents' and blocks' characteristics. Of these, 3,216 babies (90.8%) receive non-Arabic names.[8] Among parents with Arabic names, the naming decision is rather balanced since 51.1% of those parents give an Arabic name to their offspring. In contrast, among parents with non-Arabic names, the adoption of Arabic names is marginal, with a frequency of adoption of 2.8%.[9] The main difference when considering the sample of households living in public housing (reported in parentheses) is that parents with Arabic names are more likely to transmit their cultural trait to their offspring.

Since we observe overall very little adoption of Arabic names by parents with non-Arabic names in the full sample of babies from all origins, our econometric analysis will mostly look at

[8] Among them, 1,879 babies (58%) are given traditional names, that is, names that were already given in France in the early twentieth century. To identify those, we use INSEE's national database, 'Le fichier des prénoms'. Those traditional names are generally associated with Christian saint names, or names deeply ingrained in the French culture like Leo for boys or Manon for girls.

[9] The top Arabic names given by those parents are Louna for girls and Rayan for boys. Those first names are rather neutral, and are hardly selected by parents with Arabic names.

the pure transmission decision of giving an Arabic name when it is part of the original culture. To this end, we focus on households where at least one parent or grandparent is a national from Algeria, Morocco, Tunisia, the Middle East and Turkey (see Subsection 3.1 for details). The babies are thus born in France, but the parents/grandparents (babies of second/third generations, respectively) were born in an Arabic/Muslim country. However, restricting the sample to this population would leave us with a too small sample of newborn babies, especially in the public housing sector. We therefore consider children between 0 and 3 years old instead of just newborn babies to carry out this analysis. Descriptive statistics for this sample are reported in the bottom panel of Table 2. Among children with parents having themselves an Arabic name, 45% are given a non-Arabic name (658/1,447).[10] A similar pattern is observed when restricting further to households living in public housing.

2. Model and Identification of the Naming Decision

2.1. *A Simple Model of Baby Name Choice*

To estimate the channels driving the transmission of name type, we build a random utility discrete choice model of the baby naming decision. Our framework is rich enough to embed three different channels of interest (vertical, horizontal, economic) while remaining sufficiently tractable to highlight the underlying estimation issues. 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 $k(i)$, derived from choosing a given name type for its baby born in year t is defined as $U_{it}(1)$ if the name is Arabic and $U_{it}(0)$ otherwise,

$$U_{it}(\text{Baby}) \equiv V_{it}(\text{Baby}) + \epsilon_{it}(\text{Baby}),$$

where $\text{Baby} \in \{0, 1\}$ denotes alternatives, $V_{it}(\text{Baby})$ is the observed part of utility and $\epsilon_{it}(\text{Baby})$ 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 $\text{Baby}_{it} = 1$ if and only if $\Delta U_{it} \equiv U_{it}(1) - U_{it}(0) \geq 0$. Let us denote the difference in the observed part of utility as $\Delta V_{it} \equiv V_{it}(1) - V_{it}(0)$, and the difference in unobserved utility as $\varepsilon_{it} \equiv \epsilon_{it}(1) - \epsilon_{it}(0)$, such that

$$\begin{aligned}\Delta U_{it} &= \Delta V_{it} + \varepsilon_{it} \\ &= \alpha_0 + \alpha_1 \underbrace{\text{Parents}_i}_{\texttt{Vertical}} + \alpha_2 \underbrace{\mathbb{E}\left[\frac{1}{\mathcal{N}_{k(i)t}} \sum_{j \in k(i), j \neq i} \text{Baby}_{jt}\right]}_{\texttt{Horizontal}} + \alpha_3 \underbrace{\mathbb{E}[\mathcal{C}_{it}]}_{\texttt{Economic cost}} + \varepsilon_{it}, \qquad (1)\end{aligned}$$

where ΔV_{it} 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 robustness analysis). Among the $\mathcal{N}_{k(i)t}$ babies born in residential block $k(i)$ in year t, the variable Baby_{jt} codes for choices of names among babies born from other parents j

[10] 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 group of parents with non-Arabic names.

living in the block.[11] Finally, $\mathbb{E}[\mathcal{C}_{it}]$ is the *perceived* economic penalty that parents i expect to be attached to their baby if they choose an Arabic name.

The `Vertical` component captures the parental desire to transmit their own cultural type (as measured by coefficient α_1). Our specification of utility is flexible as it allows for both 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 `Horizontal` 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 parameter α_2 expected to be positive. In our data, the block k is small enough that household i is not negligible and this results in a classical Manski (1993) reflection problem. We assume that parents i form their expectations on lagged decisions of neighbours:

$$\mathbb{E}\left[\frac{1}{\mathcal{N}_{k(i)t}}\sum_{j\in k(i),\, j\neq i}\text{Baby}_{jt}\right] \equiv \frac{\sum_{\tau=1}^{\Upsilon}\sum_{j\in k(i),\, j\neq i}\text{Baby}_{jt-\tau}}{\sum_{\tau=1}^{\Upsilon}\mathcal{N}_{k(i)t-\tau}}; \tag{2}$$

that is, they expect the neighbours' current choices to be, on average, similar to those taken since year $t-\Upsilon$ (we will take $\Upsilon = 10$ in our application).

The third component (`Economic cost`) relates to economic incentives: presumably, the higher the expected penalty is, the lower the parents' desire to give their babies Arabic names. The *perceived* expected penalty 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. We now explain how it is measured in our data. Our identification strategy exploits the fact that part of the parental information set is based on information on the labour market that households retrieve from social interactions and communication with their neighbours. A straightforward approach would be to consider the unemployment differential between Arabic and non-Arabic name holders in the neighbourhood. However, the LFS is not representative at such a fine-grained level. Instead, we use the information conveyed by neighbours' occupations. The idea is that parents surrounded by neighbours working in occupations with high levels of penalty tend to update their beliefs on the extent of the penalty upwards. Formally, the *perceived* expected penalty is broken down into a block-specific informational component and an unobserved parent-specific residual component:

$$\mathbb{E}[\mathcal{C}_{it}] = \sum_{l\in\mathcal{O}}\omega_{lk(i)}\times\hat{\gamma}_l + e_{it}, \tag{3}$$

where $\mathcal{O}$ is the set of occupations, $\omega_{lk(i)}$ is the share of neighbours in block $k(i)$ working in occupation l, $\hat{\gamma}_l$ is an occupation-specific signal on the labour market penalty (see Subsection 3.1) and e_{it} is the unobserved residual parent-specific part. In the remainder of this paper, $\sum_{l\in\mathcal{O}}\omega_{lk}\times\hat{\gamma}_l$ is labelled as the *local information on penalty* (LIP) in block k. In Section A of the Online Appendix, we show how this functional form is a natural theoretical prediction in a setup where parents aim at maximising the expected utility of their child.

[11] Note that in the horizontal channel, we scale by $\mathcal{N}_{k(i)t}$ (rather than $\mathcal{N}_{k(i)t}-1$) to get a well-defined ratio in blocks where only one baby is born.

In our naming decision model, parents' utility depends directly on the expected economic cost their children face. The intensity of this relationship—captured by the parameter α_3 in equation (1)—might reflect the fact that parents are more or less likely to gather information based on their neighbours. The parameter additionally reflects the degree to which parents value their children's economic welfare. To the extent that parents discount such economic welfare heavily, this will translate into a less negative value of α_3. Accordingly, both informational sensitivity and imperfect altruism could explain parents' willingness to pay (in terms of the penalty experienced by their offspring) to perpetuate their own culture. The comparison of the coefficients α_1 and α_3 reflects the parental trade-off between their own attachment to a particular cultural type and their altruistic concern toward the future economic performance of their babies.

Combining (1), (2) and (3), utility becomes

$$\begin{aligned}\Delta U_{it} &= \Delta \mathcal{V}_{it} + \delta_{it} \\ &\equiv \alpha_0 + \alpha_1 \text{Parents}_i + \alpha_2 \frac{\sum_{\tau=1}^{\Upsilon} \sum_{j \in k(i), j \neq i} \text{Baby}_{jt-\tau}}{\sum_{\tau=1}^{\Upsilon} \mathcal{N}_{k(i)t-\tau}} + \alpha_3 \sum_{l \in \mathcal{O}} \omega_{lk(i)t} \times \hat{\gamma}_l + \delta_{it}, \qquad (4)\end{aligned}$$

where $\Delta \mathcal{V}_{it}$ is the observable utility and $\delta_{it} \equiv \alpha_3 e_{it} + \varepsilon_{it}$ is the new error term.

It is standard to specify δ_{it} as having a logistic distribution, with σ as its scaling parameter, in order to estimate the utility function (4). 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:

$$\mathbb{P}(\text{Baby}_{it} = 1) = 1/[1 + \exp(-\Delta \mathcal{V}_{it}/\sigma)]. \qquad (5)$$

The observable utility differential $\Delta \mathcal{V}_{it}/\sigma$ is retrieved from the coefficients in (4) that can be estimated readily using standard logit.

2.2. *Identification Strategy*

2.2.1. *Estimation challenges*

The key empirical challenge relates to spatial sorting of households. Before going into details, let us summarise the overall idea. In (4) and (5), a key source of identification is based on neighbours from the residential block. Neighbours are used both as: (*i*) a source of peer pressure for the horizontal transmission channel, and (*ii*) a source of information for the LIP. Since individuals tend to self-segregate, e.g., most households choose their location, our estimation could be biased by endogenous residential sorting. To address this concern, our identification strategy exploits the specificity of the French context and consists in restricting estimation to a subsample of households living in the public housing sector. Previous work (Algan *et al.*, 2016) has shown that households within public housing units are essentially exogenously allocated to their residential block, thus circumventing the issue of spatial sorting.

The horizontal transmission channel raises several estimation issues that are well known in the social interaction literature. Indeed, in (4), the realisations of $\text{Baby}_{jt-\tau}$ depend on $\Delta U_{jt-\tau}$. *Spatial sorting* might lead to a non-zero correlation between δ_{it} and $\delta_{jt-\tau}$ for households i and j belonging to the same residential block k. This would create a correlation between $\text{Baby}_{jt-\tau}$ and the error term in (4), δ_{it}, 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 devoutness 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 the 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 might also be partly driven by unobserved characteristics of the area. Our estimates could thus be biased by the endogenous spatial sorting of households. To limit this source of bias, we identify α_2 out of regressions run on a subsample of households that are allocated across the different public housing blocks within a given *département*[12] in a plausibly exogenous way (we describe the public housing allocation process at the end of this section). The combination of *département* fixed effects with quasi-random allocation of households, within a *département*, should make our econometric estimates safely immune to spatial sorting bias.

The coefficient α_3 associated with the economic cost of a name type may also be ill estimated due to self-selection into occupations and locations by parents. The methodological concern is that religious (Muslim) parents, attached to giving Arabic names to their offspring, tend to work in occupations with low discrimination, and are located in residential blocks with religious neighbours working in non-discriminating occupations. We address this issue first by controlling for parental occupation and education fixed effects. Although parental occupation is naturally not a random choice, including fixed effects for the parental occupation captures all time-invariant co-determinants of parental occupation and newborn naming choices. Second, rather than using the parental occupation as a source of information on the perceived expected penalty, we use the block-specific LIP. Thus, the remaining issue relates to 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 exogenously allocated households living in the public housing sector. We thus exploit exogenous variation in the composition of occupations across blocks in the public housing sector as a source of exogenous variation in the LIP.

Even when restricting the sample to households living in public housing, identification could be threatened by neighbourhood-level contextual drivers of both economic penalty and naming choice. For instance, housing blocks with more unskilled workers could exhibit more social discrimination against the part of the population identified with immigration from Maghreb. This could discourage parents from giving an Arabic name, *for non-economic reasons*, biasing our estimate of the economic cost effect. We address this concern by including a set of controls at the local level, such as political and anti-Islamic attitudes by occupations (aggregated at the block level), or the degree of ethnic fractionalisation. Our results show that those alternative channels are not strong determinants of naming patterns, and leave the magnitude of the economic channel estimates virtually unchanged.

A last concern is that part of the economic cost channel could operate through the horizontal channel. Indeed, our model implies that name-giving decisions by other parents in a given block should themselves be affected by the LIP. In that setting, conditioning on the local share of children with Arabic names might create a post-treatment bias for our estimate of the economic cost channel. On top of using lags in the construction of the horizontal channel variable, we address this question in the robustness Subsection 4.3 and in the Online Appendix Table E.3. Our analysis shows that the interaction between the horizontal and economic cost channels, while problematic in principle, is not quantitatively important for the economic channel point estimate.

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

2.2.2. *Public housing*

The French public housing market is very tight, and highly regulated. We provide hereafter a short overview of the allocation process of households across public housing dwellings. The most important feature for our purpose is that households have very limited control over when and where within a *département* they will be assigned if granted public housing.[13]

The main eligibility requirements for admittance into the public housing sector are to be a legal resident of France (as a French citizen or migrant with a valid residence permit) and under a certain threshold of income per unit of consumption. This income ceiling is rather high, so that around two-thirds of households living in Metropolitan France could apply for a public housing unit Jacquot (2007). The rents are also considerably lower in public housing than in private housing. As a result, there is a strong excess demand for public housing. In Paris for example, 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: family situation and household size (to ensure a suitable match with the characteristics of vacant dwellings), as well as the emergency of the application.[14]

The selection committees in charge of allocating households to vacant public housing dwellings are held at the *département* level.[15] 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 neighbourhood and in practice very few applicants (6.6%) express a preference about the area they want to be allocated within the department due to the fear of being rejected on this ground. Finally, residential mobility within the public housing sector is marginal, due to the strong shortage in the supply of public housing dwellings.

3. Estimation

In this section we start with a description of how the main explanatory variables are constructed, with a special focus on the LIP. We then proceed to our baseline estimation results, leaving our battery of robustness exercises to Section 4.

3.1. *Explanatory Variables*

The vertical transmission channel is measured by two binary variables, relating to parental characteristics relevant for the transmission of their cultural traits. The first one, *one parent/grandparent with Muslim country nationality*, codes for babies born from parents or grandparents nationals from a list of countries where Arabic names are prevalent—i.e., Algeria, Morocco, Tunisia, the

[13] Algan *et al.* (2016) provided an extensive list of tests showing the absence of self-sorting along ethnic lines across public housing dwellings in France. In Online Appendix G we provide more details on the institutional and legal aspects and descriptive statistics; we also briefly review the set of statistical tests by Algan *et al.* (2016) and expand it to several dimensions. In particular, we show that the observed allocation is not statistically different from a random process generated though Monte Carlo simulations.

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

[15] At the time of our sample, Metropolitan France was 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 *département*s 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 the municipal infrastructure.

Middle East and Turkey.[16] The second one, *one parent with Arabic name*, codes for the type of parental first names, Arabic/non-Arabic (using the same list as for babies).

The horizontal channel is measured by the *share of Arabic-named children aged 4–10 in the block* (defined in (2)). We investigate the scope of the horizontal channel by also considering naming patterns among older cohorts or among larger geographical units—*département* or *sectors* that consist of six adjacent residential blocks.

The LIP is defined as $\sum_l \omega_{lk} \times \hat{\gamma}_l$, where ω_{lk} is the share of neighbours in block k working in occupation l and $\hat{\gamma}_l$ is an occupation-specific signal on the labour market penalty attached to an Arabic name. A difficulty here relates to the abundance of ways to measure this signal. Different assumptions—in terms of labour market structure, informational frictions or parents' rationality—could point to different measurements. For instance, it is unclear whether the *true* penalty is the relevant variable to target in terms of measurement. Indeed, it would implicitly assume that parents are full-fledged econometricians in the sense that they would be able to isolate the true unemployment penalty imposed by giving an Arabic name to their baby when taking the decision. Without a clearly dominating option, we consider various measurement options of the penalty. In our baseline analysis, we use the *unconditional unemployment differential* between Arabic and non-Arabic name holders in occupation l. Denoting u_l^{a} and u_l^{na} the unemployment rate in occupation l for Arabic and non-Arabic name holders, respectively, we set $\hat{\gamma}_l = u_l^{\text{a}} - u_l^{\text{na}}$. This unconditional approach is simple and compatible with a model where agents naively attribute all the unemployment penalty they observe solely to the Arabic origin of the name. In contrast, in our first robustness exercise (Subsection 4.1), we explore the impact of measuring the signal with a *conditional* unemployment penalty retrieved from an auxiliary Mincer-type equation. This alternative approach conceptually relies on a model where sophisticated agents are able to filter out a large set of confounding factors when assessing the unemployment penalty attached to an Arabic name. Besides this demanding cognitive assumption, a caveat here is that there is no guarantee that agents use the same set of confounding factors as that used by the econometrician (e.g., non-cognitive skills). Quite remarkably, we find that the estimations based on unconditional and conditional penalties yield similar results quantitatively. Accounting for a range of observable unemployment determinants in the auxiliary Mincer equation has little effect on our coefficients of interest when estimating our main econometric equation (5). This reassuring result suggests that our findings are robust to drastically different options for measuring the labour market penalty attached to an Arabic name.

Table 3 reports basic summary statistics on unemployment rates by occupation and name type (u_l^{a} and u_l^{na}) and the associated unconditional unemployment penalty associated with Arabic names ($\hat{\gamma}_l = u_l^{\text{a}} - u_l^{\text{na}}$). For the sake of exposition (in this table only), we group together the 29 different occupations listed by INSEE into seven 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%, around three times as high as the unemployment rate of non-Arabic name holders (7%). But this average comparison hides a lot of variance across occupations. The unemployment rate of Arabic name holders among executives is only 7% and the unemployment gap with non-Arabic name holders falls to 3 points for this occupation.

[16] There is no obvious way to establish this list. We opted for countries proximate enough spatially and historically to account for a large share of the immigration flows that are relevant in terms of the vertical transmission channel. In our sample, this set of countries accounts for 80% of all parents that are nationals of a list of countries other than France, Europe, the American continent or former Indochina. The same ratio is 77% for parents living in private housing and 87% for public housing.

Table 3. *Unemployment Rates by Name Type and Occupation.*

	Unemployment rate		
	Arabic name	non-Arabic names	Unconditional penalty
Executive	0.07	0.04	0.03
Intermediate	0.14	0.05	0.09
Clerk	0.20	0.09	0.11
Blue collar (skilled)	0.20	0.07	0.13
Blue collar (unskilled)	0.29	0.15	0.14
Craftsman	0.15	0.04	0.09
Farmer	0.10	0.00	0.10
Total	0.20	0.07	0.13

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

In contrast, the unemployment rate of Arabic name holders reaches 29% among (unskilled) blue-collar workers, which represents an unemployment gap of 14 percentage points with the non-Arabic name holders in the same occupational category. The unemployment differentials for each of the 29 detailed occupations (which we use in our regressions) are presented in panel (a) of Figure E.1 in the Online Appendix. [17]

Explaining the variation in unemployment differentials across occupations goes beyond the scope of this paper. We can however think of several mechanisms that have been put forward by theoretical and empirical research that could generate this type of cross-occupation variation in discrimination on which our measurement approach relies. A type of discrimination for which there is direct empirical support in the French case is customer based: Combes *et al.* (2016) correlated the different penalty levels across occupations to how frequently employees are in direct contact with native customers. Occupations with higher levels of contact with native customers are characterised by a higher degree of employment discrimination against minority employees. Another mechanism contributing to cross-occupation variation in penalty relates to employee-based discrimination (Arrow, 1972). Applied to our context, French natives with non-Arabic names would request a higher amenity-adjusted wage to work alongside minorities with an immigration background. To the extent that prejudice/taste for discrimination among non-Arabic workers, or that the intensity of contact between workers varies across occupations, both mechanisms are expected to generate variations in unemployment penalty against Arabic names across occupations. They also suggest that we should observe less discrimination in occupations where the labour market is tight and recruitment is difficult (see Baert *et al.*, 2015 for recent evidence). Finally, empirical work has shown that stereotyping and ensuing discrimination against a given type of employee (defined based on gender, ethnicity, etc.) depends on how representative this type of employee is in a given occupation. For instance, gender stereotyping and discrimination against males has been shown to be particularly pervasive in female-dominated occupations (see, e.g., Riach and Rich, 2006 and Booth and Leigh, 2010). Along the same logic, stereotyping based on ethnicity should lead to variations across occupations in the discrimination intensity.

[17] Table F.7 in the Online Appendix presents evidence on the economic penalty at the name level by displaying the unemployment rate associated with the ten most popular non-Arabic and Arabic names.

3.1.1. *Graphical presentation of identifying variation*

As detailed in (3), the variation in the intensity of the LIP comes from nationwide differences in penalties across occupations interacted with differences in occupational shares across blocks. We present graphically these sources of variation in Figure 1, which contains two panels.

The left part of panel (a) reports the unconditional unemployment penalty for each detailed occupation. The right part of panel (a) shows the difference between the average share of each occupation in blocks belonging to the top (D10) and bottom (D1) deciles of the distribution of the LIP, computed for people living in public housing. We see that the difference between D10 and D1 is explained to a large extent by D10 blocks featuring (*i*) lower shares of public servant and employees in personal services for instance (occupations associated with medium levels of penalties), and (*ii*) higher shares of retail clerks, skilled craftsmen and low-skill manufacturing workers (associated with high levels of nationwide penalty). Taken together, these figures illustrate how the variation in the LIP comes from the co-movement between occupational shares and associated penalties.

Panel (b) of Figure 1 reports a binned scatter plot (ten bins) of the relationship between the LIP and the average share of children (aged 3 and less) with an Arabic first name. The underlying sample consists in 1,500 public housing blocks that are regrouped by decile of the distribution of LIP. Visual inspection reveals that the two measures are negatively related at the block level. As such, it constitutes a preview of our main analysis that is carried out at the individual level.

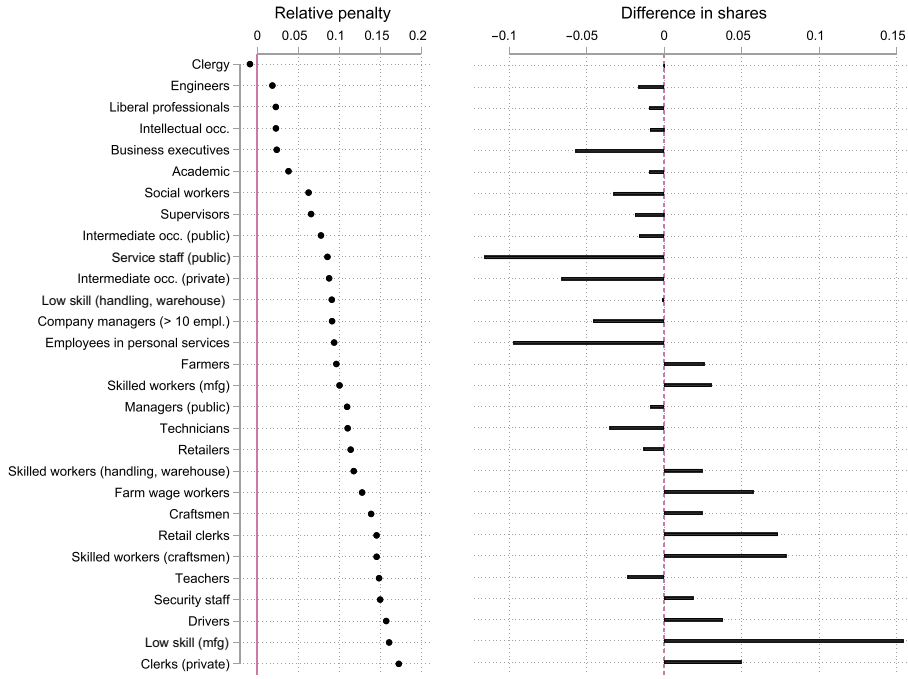
3.2. *Baseline Results*

Table 4 displays the logit estimation results of (5). The dependent variable is a binary variable coding for the Arabic origins of a baby's name. Our baseline sample consists of babies aged between 0 and 3 living in public housing. All specifications add parental occupation fixed effects, parental education fixed effects and *département* fixed effects (see Subsection 2.2). SEs are clustered at the residential block level. Average marginal effects of logit estimates are reported in all regression tables.

Columns (1) to (5) estimate regressions on babies with parents from all origins. Columns (6) to (9) restrict the sample to children born with parents or grandparents who are nationals from our list of Muslim countries described above. Those last four columns therefore condition on the first of our two vertical transmission dummy variables being turned on. The first part of the table considers the Muslim origins as a separate determinant. The second part allows for those Muslim origins to influence all determinants, and particularly how sensitive parents are with respect to the economic penalty.

The first striking result is that the coefficients on our two vertical transmission variables are positive and strongly significant in all regressions of columns (1) to (5). In the first column, we find that having a parent or a grandparent with a Muslim origin increases the probability of bearing an Arabic name by 11 percentage points. Having a parent with an Arabic name yields an even stronger effect at 28 percentage points in a sample where the baseline probability is 24%. The vertical transmission is therefore a first-order determinant, confirming the broad features of the data described in Subsection 1.2. The horizontal transmission channel also exhibits a strong positive effect, significant at the 1% threshold. With an average marginal effect at 0.11, and an SD of this variable at 0.25, the magnitude of the effect of horizontal transmission is smaller than vertical determinants, but still important.

(a) Across occupations: penalties (left), D10–D1 in block employment shares (right)



(b) LIP and the share of children (aged 3 and less) born with an Arabic first name

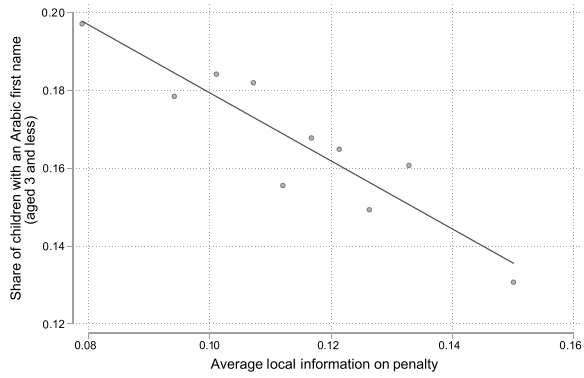


Fig. 1. *The Measure of Local Information on Penalty.*

Notes: The left part of panel (a) reports the unconditional occupational penalty in each occupation (i.e., the difference in unemployment rate between Arabic and non-Arabic name holders). The right part of panel (a) shows the difference between the average share of each occupation in blocks in the top and bottom deciles of the distribution of the LIP. The LIP is the average local occupation penalty weighted by the local share of each occupation (see (3) and the associated text). Panel (b) reports a binned scatter plot of the relationship between the average share of children aged 3 or less with an Arabic first name and the LIP. Panel (b) and the right part of panel (a) are computed for the public housing part of the sample.

Table 4. *The Choice of an Arabic Name—Baseline Results.*

Dependent variable:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Arabic name for baby								
One parent/grandp. w/ Muslim country nat.	0.11[a]	0.09[a]	0.09[a]	0.09[a]	0.09[a]				
	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)				
One parent with Arabic name	0.28[a]	0.23[a]	0.23[a]	0.23[a]	0.23[a]	0.36[a]	0.36[a]	0.36[a]	0.34[a]
	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.04)	(0.04)	(0.04)	(0.04)
Share of Arabic name in block (aged 4–10)	0.11[a]	0.09[a]	0.08[a]	0.09[a]	0.07[a]	0.03	0.06	0.07	0.00
	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)	(0.05)	(0.06)	(0.06)	(0.06)
Occupational information on penalty	−0.05								
	(0.16)								
Local unemployment penalty	−0.02								
	(0.02)								
Local information on penalty		−0.86[a]	−1.00[a]	−1.06[a]	−1.10[a]	−2.95[a]	−2.94[a]	−3.19[a]	−3.18[a]
		(0.33)	(0.34)	(0.34)	(0.36)	(1.07)	(1.09)	(1.09)	(1.08)
Local Islamophobia			0.16	0.15	0.14		0.03	0.08	−0.10
			(0.14)	(0.14)	(0.15)		(0.45)	(0.49)	(0.50)
Local ELF index			0.05	0.06	0.04		−0.15	−0.13	−0.20
			(0.04)	(0.04)	(0.05)		(0.12)	(0.12)	(0.14)
Share of Arabic name in sector (aged 4–10)				−0.02				−0.10	
				(0.04)				(0.11)	
Share of Arabic name in dept. (aged 4–10)				−0.18				0.12	
				(0.13)				(0.36)	
Share of Arabic name in block (aged 11–25)					0.01				0.08
					(0.04)				(0.10)
Share of Arabic name in block (aged 26–49)					0.04				0.15
					(0.04)				(0.11)
Share of Arabic name in block (aged 50+)					−0.05				−0.15[c]
					(0.04)				(0.08)
One parent/grandp. w/ Muslim country nat. only						Yes	Yes	Yes	Yes
Observations	2,806	3,829	3,829	3,811	3,777	992	992	987	973
Pseudo R^2	0.384	0.399	0.400	0.403	0.402	0.160	0.161	0.164	0.170
Mean probability	0.24	0.19	0.19	0.19	0.19	0.50	0.50	0.50	0.51
SD LIP		0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
SD horizontal	0.25	0.24	0.24	0.24	0.23	0.28	0.28	0.28	0.28

Notes: Logit estimates (average marginal effects). SEs, clustered at the residential block level, are in parentheses with [a], [b] and [c] respectively denoting significance at the 1%, 5% and 10% levels. All regressions include dummies for parental education level, parental occupation group, *département* of residence and years.

Column (1) presents results with a naive information structure of the economic penalty. In this specification, the two main channels of information we consider, occupation and neighbours, are entered in the regression separately. Parents are assumed to retrieve information from the (Arabic versus non-Arabic name-holders') unemployment differential (*i*) in their own occupation (nationally), (*ii*) among their neighbours—independently of their occupation. Both coefficients are negative but lack statistical significance. Regarding the self-occupation measurement, one should note that it is likely to be very noisy, since parents should not systematically infer that their children will have the same occupation as them. Besides and perhaps more importantly, our empirical strategy is not dealing with the fact that parents can self-sort into occupation, another source of bias. The second variable, unemployment penalty observed in the immediate neighbourhood, is also likely to be a very poor proxy for the perceived penalty. The Labour Force Survey is not stratified so as to be representative at the block level, introducing a considerable amount of noise in the measurement of block-level unemployment by name type.

Column (2), therefore, goes to our preferred specification of the economic cost channel, using the LIP variable. The vertical and horizontal channels keep similar magnitudes and significance levels. The coefficient associated with the LIP is negative and statistically significant at the 1% level. The estimated economic disincentive of giving an Arabic name is larger in households who live in blocks populated with individuals holding jobs most exposed to an employment penalty. How large is this effect? A natural way to proceed with quantification in this type of econometric model is to compute the predicted probability for all observations using the coefficients from column (2) twice: once using the variables at their 'true' levels, and a second time after having shocked the variable of interest. Increasing LIP by one SD (0.019) reduces the probability of giving an Arabic name by 1.59 percentage points (8.4% of the mean probability in this sample). It turns out that multiplying the average marginal effect by the SD ($-0.86 \times 0.019 = -0.01634$) yields a very reasonable approximation of the correct quantification accounting for the non-linearity of the estimator.

One might be concerned with a set of confounding factors related to the way we measure the economic cost channel. For instance, housing blocks with more unskilled workers could be more prone to social discrimination against Muslim-origin individuals, discouraging parents from giving an Arabic name *for non-economic reasons*. Another concern is that Arabic identity could be weaker in blocks with higher ethnic heterogeneity. Column (3) tries to address this concern by controlling for two local measures of potential discrimination against Arabic name holders. We first use answers to a question in a large-scale survey about attitudes toward Muslims in France, designed to be representative at the occupation level (Sauger, 2013). The most relevant question for our inquiry reads: 'Can you tell me if ISLAM means to you something very positive, fairly positive, fairly negative, or very negative?'. Following the structure of our LIP variable, we weight the occupation-specific answers by the share of each occupation in the block. As a second control, we also include a standard ELF (ethno-linguistic fractionalisation) index measuring the heterogeneity of households from different countries of origin within each block. Introducing those controls in column (3), we see that those alternative stories do not receive strong empirical support, while the impact of the economic channel remains essentially unchanged.

Columns (4) and (5) document additional features of the horizontal channel.[18] In column (4) we add the share of Arabic names for kids aged 4–10 in larger areas, either at the *sector* level or at the

[18] In our baseline analysis, we measure the horizontal channel as the share of Arabic names among 4–10 year-old children in a given block. This share is driven by the behaviour of parents irrespective of whether they have Arabic names themselves. We checked whether the strength of the horizontal channel is differentiated when this share is evaluated

département level. Those two variables measure the spatial decay of the horizontal transmission channel by looking at wider geographical units. Neither of those two variables exhibit any significant 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. Column (5) includes the share of Arabic names for older cohorts with the aim of identifying the reference group of the parents in their naming decisions. Overall, results from those two columns suggest that the horizontal channel only operates through recent choices of close neighbours, i.e., local cohorts of children under 10 years old.

In the remaining columns of Table 4, we focus on the sample of 'pure transmitters' by looking at determinants of naming decisions among babies born in France while their parents or grandparents are born with a nationality from Algeria, Morocco, Tunisia, Turkey and Middle East countries. Columns (6) to (9) replicate columns (2) to (5) on that reduced sample, yielding noticeable changes in our three channels of interest. The vertical channel (now reduced to one variable by construction of the sample) is stronger than in the general sample, while the horizontal channel becomes weaker and insignificant. Among the set of parents with Muslim/Arabic cultural background, the naming patterns of direct neighbours is much less relevant than the transmission of ones' own cultural trait. Remarkably, the effect of LIP is about three times larger for migrants from Muslim/Arabic origins than for the full sample of parents. Renewing the quantification outlined above, we find that a one SD increase in the perceived economic cost reduces the probability of giving an Arabic name to a child by around 5.5 percentage points (estimates from column (6)). This is a large effect, representing more than a 10% fall in the baseline probability of this sample (around 50%). Columns (7) to (9) add control variables for local xenophobia, peer effects from older cohorts or from larger geographic localities. Like in the full sample, the coefficients of those additional variables are not statistically significant. This finding suggests that our estimates are unlikely to be pervasively contaminated by endogenous residential sorting even for the population most likely to transmit their cultural background (in which case the coefficients for older cohorts should also be non-zero).

We explore further the heterogeneity of the economic cost channel in the Online Appendix (Table E.6). We show in that Online Appendix that the average effect of the LIP estimated in column (2) of Table 4 is entirely driven by parents with Muslim/Arabic origins. Naming decisions by other parents do not react to the economic channel. Our interpretation is that parents with a migration background have themselves probably been exposed to discrimination over their lifetimes. In contrast, other parents have not and are consequently less sensitive to the negative premium attached to Arabic names on the labour market. Accordingly, the trade-off of interest in this paper between vertical transmission and the economic cost channel seems to be relevant only for households with cultural backgrounds that make them aware and more sensitive to the economic consequences of transmitting their trait.

4. Robustness and Extensions

In this section, we perform an extensive series of robustness tests. The two most important ones are presented in detail; they respectively relate to an alternative measurement of the LIP and to the extrapolation of our findings (based on the sample of public housing tenants) to the rest of

among all parents in a block or among parents with Arabic names only. We cannot reject the hypothesis that the effects are different (see Table E.4 in the Online Appendix).

the population. Other robustness tests are reported in a more compact way—all additional details being relegated to the Online Appendix.

4.1. *Conditional Unemployment Differentials*

In our baseline empirical analysis, the LIP is based on $\hat{\gamma}_l = u_l^{\mathrm{a}} - u_l^{\mathrm{na}}$, namely the observed *unconditional* unemployment gap between Arabic and non-Arabic name holders in each occupation l. We now investigate the robustness of our findings when the LIP is based on *conditional* unemployment gaps retrieved from an auxiliary Mincer-type equation. Our aim is to test for the stability of our main coefficients when we condition in the Mincer equation with a wide range of observables (age, sex, Maghreb nationality and a set of fixed effects accounting in particular for education). Indeed, the unconditional unemployment gap might be an imperfect measure of the information truly used by the parents to assess the labour market penalty, leading to an attenuation bias in the estimation of α_3 in (4). In particular, parents might use additional information from the observed characteristics of their neighbours, such as education or country of origin, to assess the specific penalty associated with an Arabic name.

In Table 5 we replicate the set of regressions of Table 4, replacing the unconditional version of the LIP with the conditional one (omitting the first column of Table 4 that uses the naive and unsuccessful approach to measuring the economic cost channel). The analysis related to the auxiliary equations, the estimation of the conditional unemployment gaps and the construction of the LIP are detailed in Section F of the Online Appendix. Comparing the estimates obtained in each table, we observe an increase in α_3 when estimated with the second measure. Conditioning on observables therefore does not weaken (and actually strengthens) the coefficient of interest.

4.2. *How General Are Our Results?*

Beyond the usual econometric questions of identification and estimation, one can wonder how general the obtained results are. In our case, the baseline regressions are estimated using a specific part of the population, for which, we believe, biases linked to spatial sorting in particular are minimised: public housing tenants. Moreover, our central trade-off between the vertical and economic cost channels is found to be mostly relevant for households with Muslim/Arabic cultural background (see Subsection 3.2 and Online Appendix Table E.6). This group is also the main population of interest for the type of cultural transmission under study. Therefore, the most important external validity question is whether we can generalise results to all households with similar cultural backgrounds, *irrespective of whether or not they live in public housing*. We provide two exercises to that effect below.

4.2.1. *Observables in public/private housing*

Table 6 contains a number of characteristics of public versus private housing tenants among our population of interest. The list of variables covers a range of labour market outcomes and skill levels. The first point of note is contained in the last row of the table: a large fraction of households with Muslim/Arabic cultural background in France lives in public housing. In our dataset, over 50% of babies born from those households live in public housing units (992 versus 972). This implies that, independently of potential differences with the private housing sample, the sample we use for estimation and counterfactual exercises contains most of the relevant observations.

Table 5. *The Choice of an Arabic Name—Conditional Penalty.*

Dependent variable:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Arabic name for baby							
One parent/grandp. w/ Muslim country nat.	0.09[a]	0.09[a]	0.09[a]	0.09[a]				
	(0.01)	(0.01)	(0.01)	(0.01)				
One parent with Arabic name	0.23[a]	0.23[a]	0.23[a]	0.23[a]	0.36[a]	0.36[a]	0.36[a]	0.34[a]
	(0.01)	(0.01)	(0.01)	(0.01)	(0.04)	(0.04)	(0.04)	(0.04)
Share of Arabic name in block (aged 4–10)	0.08[a]	0.08[a]	0.09[a]	0.07[a]	0.04	0.06	0.07	−0.01
	(0.02)	(0.02)	(0.02)	(0.02)	(0.05)	(0.06)	(0.06)	(0.06)
Local info. on penalty (Mincer based)	−1.35[a]	−1.35[a]	−1.39[a]	−1.44[a]	−3.16[a]	−3.03[b]	−3.14[a]	−3.33[a]
	(0.37)	(0.37)	(0.37)	(0.38)	(1.17)	(1.19)	(1.18)	(1.18)
Local Islamophobia		0.18	0.16	0.14		0.02	0.04	−0.17
		(0.14)	(0.14)	(0.15)		(0.45)	(0.48)	(0.50)
Local ELF index		0.06	0.07[c]	0.05		−0.12	−0.10	−0.18
		(0.04)	(0.04)	(0.05)		(0.12)	(0.12)	(0.14)
Share of Arabic name in sector (aged 4–10)			−0.02				−0.09	
			(0.04)				(0.11)	
Share of Arabic name in dept. (aged 4–10)			−0.19				0.09	
			(0.13)				(0.36)	
Share of Arabic name in block (aged 11–25)				0.01				0.09
				(0.04)				(0.10)
Share of Arabic name in block (aged 26–49)				0.04				0.17
				(0.04)				(0.11)
Share of Arabic name in block (aged 50+)				−0.05				−0.14[c]
				(0.04)				(0.08)
One parent/grandp. w/ Muslim country nat. only	No	No	No	No	Yes	Yes	Yes	Yes
Observations	3,829	3,829	3,811	3,777	992	992	987	973
Pseudo R^2	0.401	0.401	0.404	0.403	0.160	0.161	0.164	0.170
Average prob.	0.19	0.19	0.19	0.19	0.50	0.50	0.50	0.51

Notes: Logit estimates (average marginal effects). SEs, clustered at the residential block level, are in parentheses with [a], [b] and [c] respectively denoting significance at the 1%, 5% and 10% levels. All regressions include dummies for parental education level, parental occupation group, *département* of residence and years.

Table 6. *Characteristics of Public versus Private Housing Tenants among Households with a Muslim/Arabic Cultural Background.*

	All	Public housing	Private housing	Private − public
Father is unemployed	0.17	0.20	0.15	−0.06**
Mother is unemployed	0.13	0.14	0.12	−0.02
Mother LFP	0.43	0.42	0.44	0.03
Father has high-skill occup.	0.06	0.02	0.10	0.08***
Father has middle-skill occup.	0.11	0.10	0.12	0.02
Father has low-skill occup.	0.75	0.83	0.68	−0.15***
Father: higher education	0.09	0.07	0.12	0.05***
Father's monthly wage	1,504.42	1,372.75	1,629.11	256.36***
Mother's monthly wage	1,109.38	981.20	1,214.90	233.69***
Observations	1,964	992	972	−20

Notes: This table presents the mean of a selected list of characteristics of households where at least one parent or grandparent has a nationality from a Muslim-majority country (as defined in Subsection 3.1). Within this population, it compares households living in public (column 2) versus in private housing (column 3) and tests for the statistical significance of the difference (column 4). Significance at the 1%, 5% and 10% levels is denoted by ***, ** and *, respectively.

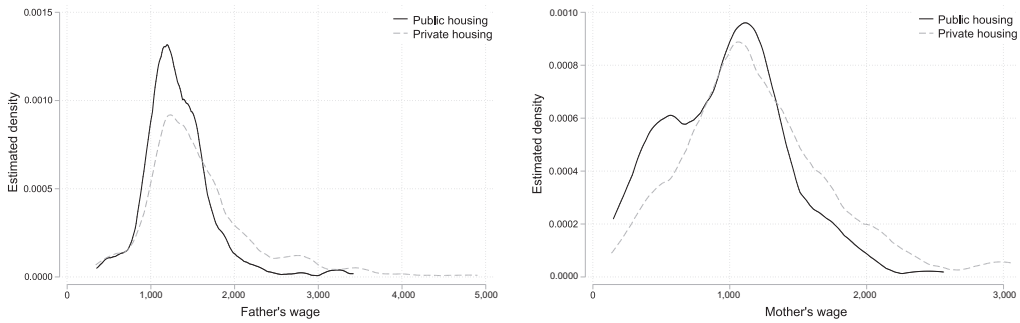


Fig. 2. *Distribution of the Monthly Wage for Fathers (left) and Mothers (right) in Public and Private Housing.*

Notes: This figure displays kernel density estimates for the distribution of public housing tenants (black solid line) and private housing tenants (grey dash line) separately among households where at least one parent or grandparent has a Muslim-majority country nationality (as defined in Subsection 3.1).

Regarding observable characteristics, Table 6 shows some expected differences in socio-economic variables: households living in public housing units have higher unemployment propensities for both parents, and the mothers' labour force participation is lower. However, the differences are not statistically significant for two out of the three variables.

Differences in terms of occupations and education between the two groups are more marked. Private housing hosts more high-skilled and much less low-skilled occupation shares of the considered population of fathers, while the middle-skill shares are fairly similar.[19] We do see significant differences in terms of average monthly wages. However, it is also instructive to look at the overall distribution of wages among the two populations, which shows substantial overlap, as illustrated in Figure 2, showing the densities of monthly wages, separately for men and women.

[19] High-skill occupations refer to executives, managers and engineers, while middle-skill occupations refer to mid-management and technicians.

The 'bump' in the low part of the distribution for women is the most noticeable difference, and is probably due to a larger share of part-time female workers in public versus private housing.

The large overlap across households in public and private housing—even in terms of observables that are key to gaining access to public housing such as monthly earnings—might seem surprising. It is probably related to several factors driving the public housing allocation process. In particular, the large weight given to the number of children in households when determining eligibility creates some overlap in overall income/education. Moreover, the well-documented public housing shortage in most areas implies long waiting lists. While the waiting time is a function of socio-economic variables, the actual access to public housing will still depend largely upon when a family first applied.

4.2.2. *Re-weighting regressions*

While the statistics displayed above are reassuring as to the extrapolation of our results to the entire population (of Muslim/Arabic background parents), small differences in terms of observable characteristics could still translate into very different behavioural responses if those characteristics exert a strong mediating effect on the channels (vertical, horizontal and economic) underpinning parental naming decisions.

In order to investigate this possibility, we carry out an additional regression-based test. An intuitive approach would consist of replicating our baseline analysis for the overall sample (or for the sample of households living in private housing). However, this would lead us back to the first-order issue of endogenous spatial sorting of households, which is likely to yield inconsistent estimates. Therefore, in keeping with our main specification estimated on the public housing sample, we follow a different route, which involves re-weighting observations of that sample. The procedure is designed so that the re-weighted public housing sample displays similar distributions for a selected number of covariates compared to the targeted sample (either the overall set of households, or those living in private housing). To that effect, we adopt the entropy balancing method of Hainmueller (2012). Those entropy weights are computed to ensure the closest possible balance between the two samples along the first moments of all the categorical variables presented in Table 6.[20] This test builds on the notion that the contrast between weighted and unweighted estimates is informative about the presence of potential heterogeneous effects in behavioural responses (Solon *et al.*, 2015). Note that the weighted and unweighted estimates could be similar for two reasons: either there is indeed limited heterogeneity along the dimension for which we are re-weighting or the weights are fairly uniformly close to 1 as there is very limited unbalance between the public housing sample and the targeted sample among the population studied.

The obtained weights can be used to check the balance between the re-weighted public housing sample and the targeted sample. Those are displayed in Table 7. We compute two sets of weights, one matching moments for the subset of households with Muslim/Arabic background living in the private housing sector (column (2)) and the other matching the overall household population (column (3)). We see that, in both cases, the balance is almost perfect, and *t*-tests (not displayed) all fail to reject any systematic differences.

We finally estimate the main specification on the re-weighted sample, using the two sets of weights contained in Table 7. Results are displayed in Table 8. Comparing columns (2) and (3), we see that, while the estimated effect of the economic channel is somewhat lower when using the weights matching the private housing sample's distribution, the difference is quite small. The

[20] We focus on categorical variables because for these matching the first moment is the natural target.

Table 7. *Re-Weighting the Public Housing Sample to Match the Distribution of All and Private Housing Samples.*

	(1)	(2)	(3)	(4)	(5)
		Public			
	No weight	Weights, private	Weights, all	Private	All
Father is unemployed	0.20	0.15	0.17	0.15	0.17
Mother is unemployed	0.14	0.12	0.13	0.12	0.13
Mother's labour force part.	0.42	0.44	0.43	0.44	0.43
Father has high-skill occup.	0.02	0.10	0.06	0.10	0.06
Father: higher education	0.07	0.12	0.09	0.12	0.09

Notes: This table presents the mean of a selected list of characteristics of households with at least one parent or grandparent a national from a Muslim-majority country (as defined in Subsection 3.1). Column (1) presents statistics for the unweighted public housing sample. Column (2) presents statistics for the weighted public housing sample where weights are computed to match the private housing sample, which is presented in column (4). Column (3) presents statistics for the public housing sample with weights computed to match the overall sample, which is presented in column (5). The weights are obtained using the entropy balancing method of Hainmueller (2012). Entropy balancing produces a set of unit weights so that the re-weighted sample satisfies a large set of pre-specified balance conditions based on known sample moments (here the first moments of binary variables).

Table 8. *Re-Weighting the Public Housing Sample to Match All/Private Housing Households.*

	(1)	(2)	(3)
		Re-weighted	
	Baseline	Private housing	All
One parent with Arabic name	0.36[a]	0.38[a]	0.37[a]
	(0.04)	(0.04)	(0.04)
Share of Arabic name in block (aged 4–10)	0.03	0.03	0.03
	(0.05)	(0.06)	(0.06)
Local information on penalty	−2.95[a]	−2.75[a]	−2.88[a]
	(1.07)	(1.07)	(1.07)
Observations	992	992	992
Pseudo R^2	0.160	0.182	0.168
Mean probability	0.51	0.51	0.51

Notes: Logit estimates (average marginal effects). SEs, clustered at the residential block level, are in parentheses with [a], [b] and [c] respectively denoting significance at the 1%, 5% and 10% levels. All regressions include dummies for parental education level, parental occupation group, *département* of residence and years. Columns (2) and (3) display results from the estimation of specifications that are identical to the baseline specification displayed in column (1) except that the samples are now re-weighted based on the entropy weights computed in order for the public housing sample to match a set of first moments of the distribution of the private housing sample (column (2)) or overall sample (column (3)). The variables used to compute the weights are presented in Table 7, namely binary variables for the father's and mother's unemployment, the mother's labour force participation, and the father's education and occupational status.

difference in estimate is even smaller when using the second set of weights—which (as expected) yields results in between the unweighted sample (column (1)) and the weights matching the private sample (column (2)). In both cases, the coefficients' stability suggests that the results obtained from the public housing sample can be plausibly extrapolated to the whole population of Muslim/Arabic background parents.[21]

[21] Naturally, our re-weighting procedure is based on a set of observable variables. We cannot exclude the possibility that unobservable characteristics between the two samples would drive the heterogeneity in the responses to the channels of interest and therefore would result in different responses among households living in private housing. Consequently,

Table 9. *The Choice of an Arabic Name—Robustness 1.*

Dependent variable:	(1)	(2)	(3)	(4)
	Arabic name for baby			
One parent with Arabic name	0.36[a]	0.35[a]	0.36[a]	0.36[a]
	(0.04)	(0.04)	(0.04)	(0.04)
Share of Arabic names in block (aged 4–10)	0.03	0.06	0.04	0.04
	(0.05)	(0.06)	(0.05)	(0.05)
Local info. on penalty	−2.95[a]	−2.23[b]	1.32	0.13
	(1.07)	(1.13)	(1.10)	(0.36)
Parent/grandp. w/ Muslim country nat. only	Yes	Yes	Yes	Yes
Specifications	Bench.	Weighted	Placebo occup.	Placebo block
Observations	992	618,374	992	986
Pseudo R^2	0.160	0.178	0.155	0.153
Mean predicted probability	0.50	0.51	0.50	0.51

Notes: Logit estimates (average marginal effects). SEs, clustered at the residential block level, are in parentheses with [a], [b] and [c] respectively denoting significance at the 1%, 5% and 10% levels. All regressions include dummies for parental education level, parental occupation group, *département* of residence and years.

4.3. *Additional Robustness Checks*

We now provide a number of additional robustness checks. Those are run on the sample of children living in public housing, aged between 0 and 3, and born from parents with Muslim/Arabic origins. This corresponds to column (5) of our baseline results in Table 4. We focus on this specification because it is both one of the most demanding in terms of sample restrictions and also the most relevant given our interest for cultural transmission decisions by immigrants and their descendants. The first set of robustness exercises is provided in the main text, while the remainder is relegated to the Online Appendix.

4.3.1. *Weighted regressions and placebo tests*

In Table 9, we perform two different exercises, both pertaining to statistical representativeness of the data. First we show that our estimation results are robust to weighting observational units by their statistical representativeness. Second, we perform two placebo tests to establish further that residential sorting is unlikely to drive our results. For the sake of comparison, we start by reporting our benchmark specification in the first column, namely the unweighted logit of column (5) in Table 4. As discussed in Section 1, the LFS is stratified at the *département* level and representativeness is thus not guaranteed at the residential block level, our level of analysis. In column (2), we report the results for a weighted logit, where the individual representativeness weights reported in the LFS are applied. We see that the 992 observed children in the estimation sample of column (1) represent 618,314 children nationwide. More importantly, we note that unweighted logit and weighted logit yield comparable coefficients. We conclude that imperfect stratification at the block level is not an issue for the estimation.

In the next two columns, we run placebo tests to rule out the possibility that our estimated LIP could be driven by some residual statistical bias attached to endogenous residential sorting. We replicate our benchmark specification on a fake sample of parents/neighbours, artificially reallocated to random occupations in column (3) and to random residential blocks in column (4). We see that, in both cases, the LIP coefficient, which is based on neighbours' occupations (see (3)), drops and also loses its statistical significance. This makes us confident that our

we choose to perform our counterfactual simulations in Section 5 solely on the basis of the public housing sample, which, as mentioned above, includes a large share of the relevant population.

Table 10. *The Choice of an Arabic Name—Robustness 2.*

Dependent variable:	(1)	(2)	(3)	(4)	(5)	(6)
	Arabic name for baby					
One parent with Arabic name	0.37[a]	0.22[c]	0.38[a]	0.33[a]	0.41[a]	0.34[a]
	(0.04)	(0.13)	(0.05)	(0.06)	(0.06)	(0.05)
Share of Arabic names in block (aged 4–10)	0.06	0.17	0.09	−0.03	0.15[c]	−0.02
	(0.06)	(0.17)	(0.07)	(0.09)	(0.08)	(0.08)
Local info. on penalty	−2.55[b]	−3.72	−1.47	−3.69[b]	−2.29	−3.98[a]
	(1.24)	(2.66)	(1.47)	(1.62)	(1.64)	(1.54)
Parent/grandp. w/ Muslim country nat. only	Yes	Yes	Yes	Yes	Yes	Yes
Specifications	Non-mixed	Mixed	2nd gen.	3rd gen.	Baby girls	Baby boys
Observations	782	143	517	432	464	470
Pseudo R^2	0.169	0.227	0.220	0.173	0.175	0.222
Mean predicted probability	0.52	0.50	0.49	0.52	0.44	0.57

Notes: Logit estimates (average marginal effects). SEs, clustered at the residential block level, are in parentheses with [a], [b] and [c] respectively denoting significance at the 1%, 5% and 10% levels. All regressions include dummies for parental education level, parental occupation group, *département* of residence and years.

identification strategy, based on the sample of children living in public housing, gets rid of endogenous residential sorting in an efficient way.

4.3.2. *Heterogeneous effects*

In Table 10, we look at heterogeneous effects by splitting the estimation sample along various relevant dimensions. In columns (1) and (2), we document the impact of cultural background by looking at couples where both parents have Arabic origins and at mixed couples, respectively. The LIP coefficient, which captures the magnitude of the economic cost channel, is larger for the latter (though less significant), while the vertical transmission motive is stronger when both parents have Arabic origins.

Columns (3) and (4) display separately the results for second generation children only (i.e., babies with parents who migrated to France) and for third generation children only (i.e., babies with parents who are born in France, but with at least one grandparent who has Muslim/Arabic origins). The economic cost channel is much larger for the latter. One interpretation is that parents with Arabic origins who are born in France are more exposed to information on discrimination. For a given level of information, the gap in the estimated effect could also be driven by different preferences, with newly arrived migrants displaying a lower willingness to adjust the vertical transmission of their culture in response to signals about the likely economic cost of such transmission for the economic well-being of their child. Parents who are first-generation migrants appear to put less weight on the economic cost of their naming decisions and attach somewhat more importance to the vertical channel. First-generation parents may be less able to use their surroundings to gather information on the labour market or less aware of the functioning of the labour market when they have kids. The parameter additionally reflects the degree to which parents value the economic welfare of their children. To the extent that parents discount such economic welfare heavily, this will translate into a less negative value of α_3. Accordingly, either differences in the discount rate attached to their children's economic welfare or differences in information sensitivity could explain first-generation parents' higher willingness to pay—in terms of the penalty experienced by their offspring—to perpetuate their own culture.

In columns (5) and (6) of Table 10, we study how naming decision determinants differ across genders by splitting the sample in two: baby girls and baby boys, respectively. We see that the

horizontal transmission channel is significant for girls, but not for boys. The reverse is true of the information on economic penalty whose effect is significantly negative for boys only. Previous works by sociologists who study naming patterns among minorities have documented that parents are more open to 'creative names' for girls than for boys, who tend to receive more traditional names (Sue and Telles, 2007; Gerhards and Silke, 2009). This difference is likely to result in a lower rate of name convergence for boys than for girls. This is reflected in the mean predicted probability that girls receive an Arabic name which is much lower (0.44) than that of boys (0.57). Interestingly, this lower rate of assimilation for boys occurs despite the fact that the marginal effect of the economic channel (while negative for both genders) is much stronger for boys than for girls (which has a *p*-value of 0.101). Finally, we see that the horizontal channel is stronger for girls while the vertical channel is of similar magnitude across genders. Overall, our results suggest that parents are either more cognizant or more sensitive to the economic penalty imposed upon their baby boys than baby girls. This is consistent with parents' envisioning traditional gender roles for their children in the labour market and being more sensitive to peer effects when choosing girl names.

4.3.3. *Intensity of vertical transmission*

Our measure of the vertical transmission channel is a binary variable equal to 1 whenever one of the parents has an Arabic first name. We differentiate the effect depending on whether the mother only, the father only or both parents have an Arabic first name. Households where none of the parents have an Arabic name is the reference group. Results are presented in Online Appendix Table E.5, in which we see that couples where the mother has an Arabic name are 19% more likely to give an Arabic name to their child than couples where neither of the parents carries an Arabic name. Interestingly, the father having an Arabic name is associated with a somewhat larger increase in the probability (+22%). Finally, we see that having both parents carry an Arabic name is associated with a substantially larger effect (+26%). The finding that the father's cultural marker has a larger influence on naming is consistent with previous sociological studies (see Sue and Telles, 2007, who studied naming patterns among Hispanics in the United States). Once we focus on individuals with an immigration background from Muslim countries, we see the same ordering in terms of the size of the vertical channel across categories of couples.

4.3.4. *Use of measure of relative penalty rather than in levels*

Our analysis is based on a measure of occupation-specific penalties expressed in *levels*, i.e., the percentage point differences in unemployment rates between Arabic and non-Arabic name holders. In Section E.1 of the Online Appendix, we show that, using *relative* unemployment penalties, i.e., the ratio of unemployment rates of Arabic name holders versus that of individuals with non-Arabic names, leads to quantitatively similar results (see in particular Table E.2 of that Online Appendix). We further show that our results are unchanged when using the measure of penalty in levels and controlling for the baseline unemployment rate, i.e., unemployment among non-Arabic-name workers, as predicted by the occupations of neighbours (see Table E.1 in particular).

4.3.5. *Interaction between the horizontal and economic cost channels*

A concern that might arise regarding the economic channel is that part of it could operate through the horizontal channel. Indeed, the intensity of Arabic name giving in a given block is the result of individual decisions that are themselves a function of the LIP (since the naming decision

is also the LHS variable, this results in a potential manifestation of the reflection problem). In that setting, conditioning on the share of children with Arabic names might bias our estimate of the marginal effect of the LIP due to post-treatment bias (Imai *et al.*, 2011). A feature of our benchmark regressions, which should mitigate the possibility that the horizontal channel is a product of the LIP, is that we measure the horizontal channel as the *lagged* block-level share of Arabic names given to children. There is thus a minimum of 1 year between the naming decision of households we are analysing (as the LHS variable) and the realisation of the same outcome among peers in our measure of the horizontal channel (as a RHS variable).

In order to further assess how the magnitude of our economic channel estimates depends upon the presence of the horizontal channel, we provide robustness regressions, where we re-estimate our baseline specification without measuring the horizontal channel. The results, reported in Table E.3 of the Online Appendix, show that omitting the horizontal channel variable leaves the estimated coefficient associated with the economic cost virtually unchanged. This suggests that, while interactions between the horizontal and economic cost channels could, in principle, be an issue for our estimation, they do not seem to matter in practice.

5. Quantification and Welfare Analysis

We now turn to quantifying the effects of vertical, horizontal and economic channels in the naming decision. We first analyse the short-run contributions of each channel. We then perform a welfare analysis. In Section D of the Online Appendix, we also quantify the long-run effects taking into account the dynamics of inter-generational cultural transmission. All the analysis is based on estimates from our baseline Table 4.

5.1. *Short-Run Effects*

In Table 4, coefficients are reported as average marginal effects over choices in our sample. Therefore, the change in the baseline probability of an independent change in each channel is easy to interpret (see Subsection 3.2). An alternative, and interesting, 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 adopt the following strategy: we start by running our benchmark regression to estimate the coefficients of interest, which gives us the benchmark probability of transmitting an Arabic first name in the sample. Then we run the counterfactual by changing the values for one or more explanatory variables. For instance, we shut down the economic cost channel by forcing the LIP variable to be zero for the whole sample. The logit formula (5) provides the counterfactual naming probability for each observation. Summing those over the sample gives the counterfactual number of babies born with an Arabic name in each experiment. This procedure ensures that the probability remains within the admissible range while doing a 'what if' experiment.

Results are reported in Table 11, where different lines present different scenarios. We focus on the sample of babies born to parents with a Muslim/Arabic background and living in public housing (the point estimates of column (6) in Table 4). The first line reports the true number of babies with an Arabic name, 501 in this sample; they represent 320,851 babies nationally when survey weights are applied. The second line is the number of babies born with Arabic names as predicted by the benchmark regression of column (6). We then remove (in the third row) the vertical channel associated with the parental name. The predicted number of Arabic naming

Table 11. *Quantification of the Three Channels.*

	# babies with Arabic name		Mean Δ welfare w.r.t. benchmark	
Scenario:	Count	Weighted count	Change	Weighted change
True figure	501	320,851	–	–
Benchmark	501	316,292	–	–
No vertical (parental name)	221	139,636	−0.555	−0.567
No horizontal	491	310,516	−0.026	−0.024
No penalty	783	490,945	1.098	1.109
No ghetto	500	316,275	0.002	0.005
No foreign names	0	0	−0.822	−0.838

Notes: This table uses logit estimates (column (6) of Table 5) based on the sample of 992 babies (0–3 years old) with Arabic origins and living in public housing (representing 618,374 nationally). Each line presents a scenario, removing in turn one of the channels of influence in the regression.

decisions falls to 221 in that case, that is, 44% of true births. This is a quite drastic cut, especially when compared to the horizontal channel, where a similar thought experiment removes only 2% of Arabic naming decisions from the benchmark. The economic channel has a much stronger effect than the horizontal one: removing the economic penalty completely increases the number of babies receiving an Arabic name by 56%. The line 'no ghetto' shows the results from a slightly different experiment. In this scenario, all blocks in the country had the same neighbourhood composition and the same information on unemployment penalty. This amounts to considering the predicted number of babies when averaging the horizontal and penalty variables, which induces effects that almost cancel out in naming choices on average.

5.2. *Welfare Analysis*

Our random-utility, discrete-choice model of naming choice allows for a quite simple characterisation of welfare changes associated with the different thought experiments of Table 11. A natural metric for welfare in this model is the expected value of *parental* maximum utility between naming choices (the theoretical analysis of welfare analysis in our model is provided in Online Appendix B). This varies across households, and we average this welfare over our sample (a simple average in the third column of Table 11, and a weighted one in the fourth). Note that we consider welfare from the parental perspective—a natural approach in this short-run analysis, where parents are the decision makers in term of naming choice. When it turns to modelling inter-generational dynamics (Online Appendix D), the approach has to be qualified because our underlying model of cultural transmission is based on an assumption of imperfect altruism (see our theoretical discussion in Subsection 2.1): this implies that children's welfare and parents' welfare cannot be treated as one and the same.

The absolute level of welfare has no meaningful unit in the logit model—as noted by Anderson *et al.* (1992) and Train (2009)—and a natural way to quantify welfare changes is first to take the difference between welfare in each scenario and the benchmark case and then compare across scenarios. Looking at the last two columns of Table 11, we see, for instance, that the negative impact of removing the vertical transmission motive would be more than twenty times larger than the effect of removing the horizontal one. Considering economic penalty, the effect naturally varies according to the cut in the additional unemployment rate associated with Arabic names. If this penalty were brought to zero, the gain in welfare would be about twice as large as that arising from the utility boost linked to vertical transmission. Figure 3 spans over a

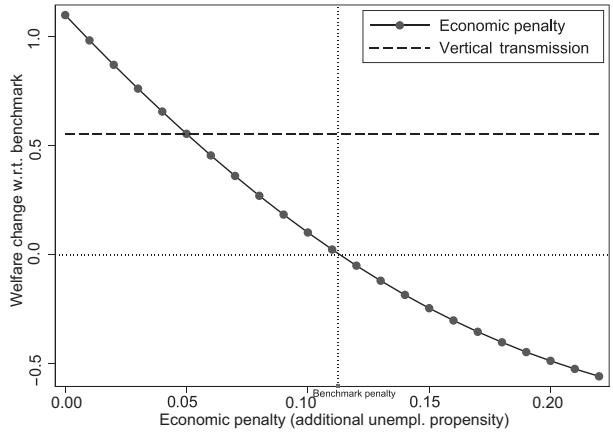


Fig. 3. *Welfare in the Short-Run, Economic Penalty and Vertical Transmission.*

wider set of changes in economic penalty and compares it to the welfare changes associated with vertical transmission. The x axis reports the counterfactual economic penalty (percentage point differences in unemployment rates). The y axis measures welfare change differences with the benchmark level (with zero change occurring at the sample average of economic penalty, around an 11 percentage point difference). We also represent the utility gain (with respect to the benchmark) associated with vertical transmission. An interesting conclusion from this figure is that we are now able to gauge cultural attachment strength in monetary units. Indeed, the vertical transmission motive of one's cultural trait is equivalent in terms of welfare gains to a cut by around half the perceived economic penalty associated with that trait. Since, in this sample, this amounts to cutting the penalty by around 6 percentage points, using the estimates of unemployment-related income loss reported in Section F of the Online Appendix, the vertical transmission channel is found to be on the same indifference curve as a 3% upward shift in lifetime income of one's child.

It is possible to look for differences in the monetary valuation of cultural attachment across households. Probably the most meaningful source of heterogeneity is related to whether the parents choosing a name for their baby are themselves first-generation migrants from Arabic countries versus being born in France from migrating Arabic parents. This distinction is done in columns (3) and (4) of Table 10, and we reproduce in Online Appendix C the equivalent of Figure 3 for the sub-samples of first and second generation parents separately (Figure C.1). Interestingly, we find a much stronger monetary equivalent of the vertical transmission motive for first-generation migrants. For those, the vertical transmission channel is on the same indifference curve as a 6.2% upward shift in lifetime income of one's child while for second generation parents this figure amounts to a 2.3% upward shift in lifetime income.

Finally, we consider an experiment where France would return to historical naming regulations. Between 1803 and 1993, the choice of first names was essentially restricted to Saints' names, names from ancient Greece and Rome and names from the Bible. The legal procedure was that a civil officer had to state whether the name proposed by the parents respected the 1803 Napoleonic law. If the answer was negative, the parents had to challenge the decision in court. Foreign names were hardly tolerated at all before a 1987 revision explicitly asked civil officers to be more liberal

with names coming from a 'foreign or French tradition, whether national or local'. Note that the computation of this scenario involves shutting down not only all three channels emphasised in our paper, but also the occupation, regional and educational controls we have in the regression, in order to generate a predicted number of babies with Arabic names of 0. The mean welfare loss from this return to a strict ban on foreign—and therefore Arabic—names would be substantial, around 50% larger than the cut of the vertical channel alone.

5.2.1. *Long-run implications*

In Section D of the Online Appendix, we explore the long-run implications of our structural model on naming patterns. To this purpose, we consider a simple extension of the static model described in Section 2 that accounts for inter-generational cultural transmission dynamics. We restrict our focus to a partial equilibrium setup where any potential feedback effect of naming patterns on the economic penalty is ignored—admittedly an important simplification in a long-run perspective. We come up with two main findings. First, our quantification shows that the long-run share of Arabic name holders predicted by our structural model for the population of Muslim/Arabic background individuals living in public housing should converge to 9%, which is much smaller than the actual one in our sample (48%). Hence, the actual share is still far from its steady-state value and the transitory dynamics is expected to bring it down in the future. This feature might be explained by the fact that migration from Arabic countries is still a quite recent phenomenon in France, and most babies born in the 2003–2007 period belong to the third generation of migrants only. Second, our analysis confirms that the economic cost channel is also a key driver of cultural transmission in the long run. In a counterfactual scenario, where the labour-market penalty attached to Arabic names is artificially brought to zero, the long-run share of Arabic name holders should converge to a much larger steady-state level, namely 36% instead of 9%. Finally, we refrain from drawing any strong conclusion with respect to long-run welfare effects. As discussed above, defining a welfare criterion in an inter-generational model of cultural transmission would require taking a stance on the degree of parental altruism. This would imply extending the model beyond the scope of current paper.

6. Conclusion

While it might seem natural to consider culture as a deep individual characteristic, our paper shows that the cultural choices made by a person cannot be completely insulated from the economic context in which he/she operates. We focus on one cultural trait that has the advantage of being easily measurable, identified by social sciences as a key marker of cultural identity and has economic consequences: first names.

Our results show that the information about economic factors available to parents deeply shape individual decisions of cultural transmission. While the vertical channel plays a key role in the cultural transmission process, parents do account for the information about the economic cost of their cultural trait in their naming decisions. Counterfactually reducing the economic penalty on Arabic names to zero, the annual number of babies born with an Arabic name in France would be more than 50% larger. The horizontal channel, which has been the focus of much attention in the social interaction literature, is found to be much less important in our case. Our theory-based estimates allow us to perform a welfare analysis where we gauge cultural attachment strength in monetary units. We find that allowing for a vertical transmission channel provides the same shift to parents' utility as a 3% raise in lifetime income of the child. We also show that a return to an

old regulation banning choice of names of foreign origins would cause very important losses to the well-being of parents.

While we have focused on naming decisions, our paper opens new questions on the welfare effects of public policies aiming at promoting or restraining expressions of cultural identity, such as wearing religious signs in public areas. It also raises questions about the use in academic papers of cultural traits as determinants of economic outcomes. It seems clear from our results that at least some aspects of culture cannot be considered as exogenous to what happens in the economic sphere.

Sciences Po, France & CEPR, UK
Institut des Politiques Publiques (IPP) & Paris School of Economics (PSE), France & IZA, Germany
Sciences Po, CEPII, France & CEPR, UK
University of Lausanne, Switzerland & CEPR, UK

Additional Supporting Information may be found in the online version of this article:

Online Appendix
Replication Package

References

Abramitzky, R., Boustan, L.P. and Eriksson, K. (2016). 'Cultural assimilation during the age of mass migration', Working Paper, National Bureau of Economic Research.
Abramitzky, R., Boustan, L. and Eriksson, K. (2020). 'Do immigrants assimilate more slowly today than in the past?', *American Economic Review: Insights*, vol. 2(1), pp. 125–41.
Adida, C.L., Laitin, D.D. and Valfort, M.A. (2010). 'Identifying barriers to Muslim integration in France', *Proceedings of the National Academy of Sciences*, vol. 107(52), pp. 22384–90.
Akerlof, G.A. and Kranton, R.E. (2000). 'Economics and identity', *The Quarterly Journal of Economics*, vol. 115(3), pp. 715–53.
Alesina, A. and Giuliano, P. (2015). 'Culture and institutions', *Journal of Economic Literature*, vol. 53(4), pp. 898–944.
Algan, Y. and Cahuc, P. (2013). 'Trust and growth', *Annual Review of Economics*, vol. 5(1), pp. 521–49.
Algan, Y., Dustmann, C., Glitz, A. and Manning, A. (2010). 'The economic situation of first and second-generation immigrants in France, Germany and the United Kingdom', Economic Journal, vol. 120(542), pp. F4–30.
Algan, Y., Hémet, C. and Laitin, D.D. (2016). 'The social effects of ethnic diversity at the local level: A natural experiment with exogenous residential allocation', *Journal of Political Economy*, vol. 124(3), pp. 696–733.
Anderson, S.P., De Palma, A. and Thisse, J.F. (1992). *Discrete Choice Theory of Product Differentiation*, Cambridge, Massachussets: MIT press.
Arrow, K.J. (1972). 'Models of job discrimination', in (A.H. Pascal, ed.), *Racial Discrimination in Economic Life*, pp. 83–102, Lexington: D. C. Heath.
Atkin, D., Colson-Sihra, E. and Shayo, M. (2019). 'How do we choose our identity? A revealed preference approach using food consumption', Working Paper, National Bureau of Economic Research.
Baert, S., Cockx, B., Gheyle, N. and Vandamme, C. (2015). 'Is there less discrimination in occupations where recruitment is difficult?', *ILR Review*, vol. 68(3), pp. 467–500.
Bertrand, M. and Mullainathan, S. (2004). 'Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination', *American Economic Review*, vol. 94(4), pp. 991–1013.
Bisin, A., Patacchini, E., Verdier, T. and Zenou, Y. (2011). 'Formation and persistence of oppositional identities', *European Economic Review*, vol. 55(8), pp. 1046–71.
Bisin, A., Patacchini, E., Verdier, T. and Zenou, Y. (2016). 'Bend it like Beckham: Ethnic identity and integration', *European Economic Review*, vol. 90, pp. 146–64.
Bisin, A., Topa, G. and Verdier, T. (2004). 'Religious intermarriage and socialization in the United States', *Journal of Political Economy*, vol. 112(3), pp. 615–64.
Bisin, A. and Verdier, T. (2001). 'The economics of cultural transmission and the dynamics of preferences', *Journal of Economic Theory*, vol. 97(2), pp. 298–319.
Bisin, A. and Verdier, T. (2011). 'The economics of cultural transmission and socialization', in (J. Benhabib, A. Bisin and M. Jackson, eds.), *Handbook of Social Economics*, pp. 339–416, Amsterdam: North-Holland.

Booth, A. and Leigh, A. (2010). 'Do employers discriminate by gender? A field experiment in female-dominated occupations', *Economics Letters*, vol. 107(2), pp. 236–8.
Combes, P.P., Decreuse, B., Laouenan, M. and Trannoy, A. (2016). 'Customer discrimination and employment outcomes: Theory and evidence from the French labor market', *Journal of Labor Economics*, vol. 34(1), pp. 107–60.
Duguet, E., Leandri, N., L'horty, Y. and Petit, P. (2010). 'Are young French jobseekers of ethnic immigrant origin discriminated against? A controlled experiment in the Paris area', *Annals of Economics and Statistics/Annales d'Économie et de Statistique*, vol. 99/100, pp. 187–215.
Fernández, R. (2011). 'Does culture matter?', in (J. Benhabib, A. Bisin and M. Jackson, eds.), *Handbook of Social Economics*, pp. 481–510, Amsterdam: North-Holland.
Fouka, V. (2019). 'How do immigrants respond to discrimination? The case of Germans in the US during World War I', *American Political Science Review*, vol. 113(2), pp. 405–22.
Fouka, V. (2020). 'Backlash: The unintended effects of language prohibition in US schools after World War I', *The Review of Economic Studies*, vol. 87(1), pp. 204–39.
Fryer, J.R.G. and Levitt, S.D. (2004). 'The causes and consequences of distinctively Black names', *The Quarterly Journal of Economics*, vol. 119(3), pp. 767–805.
Gerhards, J. and Silke, H. (2009). 'From Hasan to Herbert: Name-giving patterns of immigrant parents between acculturation and ethnic maintenance', *American Journal of Sociology*, vol. 114(4), pp. 1102–28.
Goel, D. and Lang, K. (2019). 'Social ties and the job search of recent immigrants', *ILR Review*, vol. 72(2), pp. 355–81.
Goux, D. and Maurin, E. (2007). 'Close neighbours matter: Neighbourhood effects on early performance at school', Economic Journal, vol. 117(523), pp. 1193–215.
Guiso, L., Sapienza, P. and Zingales, L. (2008). 'Social capital as good culture', *Journal of the European Economic Association*, vol. 6(2–3), pp. 295–320.
Guiso, L., Sapienza, P. and Zingales, L. (2016). 'Long-term persistence', *Journal of the European Economic Association*, vol. 14(6), pp. 1401–36.
Hainmueller, J. (2012). 'Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies', *Political Analysis*, vol. 20(1), pp. 25–46.
Hainmueller, J. and Hiscox, M.J. (2010). 'Attitudes toward highly skilled and low-skilled immigration: Evidence from a survey experiment', *American Political Science Review*, vol. 104(1), pp. 61–84.
Hainmueller, J. and Hopkins, D.J. (2014). 'Public attitudes toward immigration', *Annual Review of Political Science*, vol. 17, pp. 225–49.
Heckman, J.J. (1998). 'Detecting discrimination', *Journal of Economic Perspectives*, vol. 12(2), pp. 101–16.
Hellerstein, J.K., McInerney, M. and Neumark, D. (2011). 'Neighbors and coworkers: The importance of residential labor market networks', *Journal of Labor Economics*, vol. 29(4), pp. 659–95.
Imai, K., Keele, L., Tingley, D. and Yamamoto, T. (2011). 'Unpacking the black box of causality: Learning about causal mechanisms from experimental and observational studies', *American Political Science Review*, vol. 105(4), pp. 765–89.
INSEE. (2016). *France, Portrait Social*, INSEE.
Jacquot, A. (2007). 'L'occupation du parc HLM: un éclairage à partir des enquêtes logement de l'INSEE', Document de Travail F0708, INSEE.
Jouniaux, L. (2001). *Les 20,000 plus beaux prénoms du monde*, Paris: Hachette Pratique.
Lieberson, S. (2000). *A Matter of Taste: How Names, Fashions, and Culture Change*, New Haven: Yale University Press.
Lleras-Muney, A. and Shertzer, A. (2015). 'Did the Americanization movement succeed? An evaluation of the effect of English-only and compulsory schooling laws on immigrants', *American Economic Journal: Economic Policy*, vol. 7(3), pp. 258–90.
Manski, C.F. (1993). 'Identification of endogenous social effects: The reflection problem', *The Review of Economic Studies*, vol. 60(3), pp. 531–42.
Maurin, E. and Moschion, J. (2009). 'The social multiplier and labor market participation of mothers', *American Economic Journal: Applied Economics*, vol. 1(1), pp. 251–72.
Mazumder, S. (2019). 'Becoming white: How mass warfare turned immigrants into Americans', Working Paper, available at : SocArXiv agjsm, Center for Open Science.
Nunn, N. and Wantchekon, L. (2011). 'The slave trade and the origins of mistrust in Africa', *American Economic Review*, vol. 101(7), pp. 3221–52.
Riach, P.A. and Rich, J. (2006). 'An experimental investigation of sexual discrimination in hiring in the English labor market', *The B.E. Journal of Economic Analysis & Policy*, vol. 6(2), pp. 1416.
Rozenas, A. and Zhukov, Y. (2019). 'Mass repression and political loyalty: Evidence from Stalin's 'terror by hunger'', *American Political Science Review*, vol. 113(2), pp. 569–83.
Sauger, N. (2013). 'Citoyens et candidats en 2012: Convergence ou divergence de vues? Enquête TNS SOFRES/TRIELEC 2012', in (É. Lecerf and O. Duhamel, eds.), *L'état de l'opinion 2013*, pp. 99–116, Paris: Editions Seuil.
Shayo, M. (2009). 'A model of social identity with an application to political economy: Nation, class, and redistribution', *American Political Science Review*, vol. 103(2), pp. 147–74.
Solon, G., Haider, S.J. and Wooldridge, J.M. (2015). 'What are we weighting for?', *Journal of Human Resources*, vol. 50(2), pp. 301–16.

Sue, C.A. and Telles, E.E. (2007). 'Assimilation and gender in naming', *American Journal of Sociology*, vol. 112(5), pp. 1383–415.
Tabellini, G. (2008). 'The scope of cooperation: Values and incentives', *The Quarterly Journal of Economics*, vol. 123(3), pp. 905–50.
Train, K.E. (2009). *Discrete Choice Methods with Simulation*, Cambridge: Cambridge University Press.
Voigtländer, N. and Voth, H.J. (2012). 'Persecution perpetuated: The medieval origins of anti-semitic violence in Nazi Germany', *The Quarterly Journal of Economics*, vol. 127(3), pp. 1339–92.