Cambridge Working Papers in Economics

Does identity matter?

Zs Koczan

Does identity matter?

Zs. Koczan[*]

King's College, King's Parade, CB2 1ST, University of Cambridge, Cambridge, UK

Abstract

We examine the question of whether identity is just a 'label' or whether it matters in affecting outcomes, such as education, employment or political orientation, using data on Turkish and ex-Yugoslavian second generation immigrants in Austria and Germany. We begin with an empirical investigation of identity formation, with a focus on parental investment in their child's identity, and use this to understand the impact of the child's own identity on own outcomes, a generation later. The results suggest that identity does not have a significant effect on education, employment and political orientation, thus suggesting that a strong ethnic/ religious minority identity does not constrain the second generation or hamper socioeconomic integration.

Keywords: Identity, second generation immigrants, integration
JEL classification: F22, J15, O15

1. Introduction

Identity is generally interpreted as a self-definition, a 'narrative' that people tell themselves and others, as the answer that they give to the question 'who am I?'. As such, it is a well-defined concept in the psychology literature that can be measured using survey questions of the type 'to what extent do you feel...?'. The literature suggests that identity is formed in childhood and is likely to be shaped by various events in a person's life, hence does evolve

over time. While it is an inherently multidimensional concept, including gender and professional identities, to name only a few, the focus here is on ethnic or national feelings of belonging[1].

Recent popular debates suggest that interest in whether national identities matter has grown, while the literature in economics has remained largely theoretical (Akerlof and Kranton, 2000). In this paper, we aim to bolster the empirical evidence on the question of whether identity matters for socio-economic integration. Is identity just a 'label', or does it affect outcomes, such as education, employment or political orientation?

We use data on Turkish and ex-Yugoslavian second generation immigrants (the children of immigrant parentage born in the country of migration) in the former guest worker recruiting countries, Austria and Germany. We begin with an intergenerational model to examine the roles of parental investments (including the choice of language in which to raise their children), personal characteristics, and peer effects and environmental influences in the development of the identity of the second generation[2]. In the next stage, we examine the effects of identity on education, employment and political orientation. A particular concern is the potential endogeneity of identity: we use those parental investments in early childhood that affect identity but can plausibly be thought of as not affecting current outcomes to deal with these issues. The main focus will be on the language in which the child was raised, controlling for fluency in the majority language but we also examine the robustness of the estimates to alternative sets of instruments.

Language choice has always been closely tied to the notion of identity and underpins popular notions of identity formation and integration. As Turkey's prime minister, Recep Tayyip Erdoğan visited Germany in February 2011, his statement 'our children must learn German, but they must learn Turkish first' in a speech in Düsseldorf sparked a large social debate in an

---

[1]These are sometimes referred to in the literature as ethnic self-identifications, for a detailed discussion please see section 3.2.

[2]For simplicity and due to data constraints identity investments are treated here as a myopic problem, i.e. it is assumed that parents invest in the identity of their children without considering the economic consequences for the children in say 20 years. While parents could invest in the identity of their children for a variety of reasons including economic as well as 'cultural' motives (e.g. they could care about identity for its own sake), their underlying motivations do not play a role in the following analysis as we focus on the children and treat parental investments as given from their perspective.

atmosphere where Germany's chancellor, Angela Merkel declared in 2010 that multiculturalism in Germany had 'utterly failed'. Given the resulting controversy, we may believe that it was not only about the extra resources required in kindergarten to ensure equal chances, but also about a possible implicit assumption that (first) language may have a strong link with future identity and socio-economic integration.

We hope to contribute to the literature on identity by beginning with an empirical investigation of identity formation, with a focus on parental investments in their child's identity and then using these insights to overcome the endogeneity problems that often arise in regressions attempting to measure the effect of identity on economic (or political) outcomes. We explicitly allow for both single and multiple identities.

The paper is structured as follows. Section 2 discusses the current literature on identity and (second generation) immigrants. Section 3 introduces the data used. Section 4 outlines the theoretical basis for the first stage, the identity formation mechanism, discussing the model of Bisin et al. (2006) and the second stage, the question whether identity matters for economic and political outcomes, contrasting standard economic arguments with the recent work of Akerlof and Kranton (2000, 2002, 2005, 2010). Relying on this theoretical framework, section 5 then outlines our empirical methodology and presents the results, including robustness tests. Section 6 concludes.

2. Literature review

The paper is motivated by two key issues raised by the current literature: first, that there is little empirical evidence on the factors that affect identity formation and second, that equally, there is little evidence of the impact of identity on socio-economic outcomes. This paper hopes to help fill this gap.

Numerous papers examine the identity of (second generation) immigrants relying on qualitative analysis, especially semi-structured interviews, based on small samples with few cross-country or cross-group comparisons (see e.g. Waters, 1994; Portes and MacLeod, 1996; Zephir, 2001; MacFadden, 2004; Zimmerman, Zimmerman and Constant, 2006; Lewandowska, 2007; Clark, 2008; Somerville, 2008). Other papers (Manning and Roy, 2007; Georgiadis and Manning, 2009) examine models of identity formation to explain choices of identity and compare the utility obtained from different choices of group membership. Our approach is related to these models but aims to extend this framework by then linking identity to economic and political outcomes.

Few papers link identity to economic and political outcomes. Although numerous European studies dealt with the education and employment of second generation immigrants, these generally only looked at the effect of ethnicity rather than identity, examining the question whether children of immigrants inherit their parents' low socio-economic positions (e.g. Crul and Schneider, 2009; Adsera and Chiswick, 2004; Heath and Cheung, 2007; De Coulon and Wadsworth, 2008).

The impact of identity on outcomes is conflated by the endogeneity of identity as models of identity formation make plain. Battu and Zenou (2010) estimate the effect of identity on employment, instrumenting identity with variables measuring whether individuals have experienced racial harassment, if their parents made the decision in choosing their wife or husband and if they prefer a school of their own religion for their children. Unfortunately the measured effect could be confounded by reverse causality. Nekby and Rodin (2007) look at the consequences of identity for labour market outcomes in Sweden, while Pendakur and Pendakur (2005) look at the relationship between ethnic minority identity and the use of informal networks in finding a job, but both papers wrestle with the endogeneity of residential location.

Casey and Dustmann (2010) offer the only analysis of identity formation and the consequent effects of identity on labour market outcomes. They develop a model of parental identity investments in which if there is no earnings disadvantage from a minority identity then it is optimal for the parent if the child's identity is equal to the parent's identity but parents might 'restrict' minority identity investments to the extent that there is some disadvantage. They acknowledge that their findings cannot be interpreted as causal, but argue that if economically successful individuals feel more strongly German then their estimate of identity is an upper bound, and that if this effect is symmetric this is bounded below by the coefficient on minority identity[3]. This reasoning assumes a mutually exclusive relationship between minority and majority identity, which (as will be discussed in detail in section 4.1) is not necessarily the case.

A closely linked study is also that by Schüller (2011), which evaluates the effect of parents' ethnic identity on the educational attainments of second generation immigrants in Germany using the German Socio-Economic Panel and finds that both minority and majority identities have a positive

---

[3]They also note that a simultaneity bias would act in the same direction.

4

significant effect on the child's probability to be placed in a higher secondary schooling track. This paper argues that the positive impact of majority identity works exclusively through mothers, while the impact of minority identity is specific to fathers. However, results should be interpreted with caution as they may be affected by omitted variable bias - in particular, specifications do not control for child's identity.

We offer an alternative attempt to parse the effect of the child's identity on outcomes where individuals are allowed to maintain multiple or overlapping identities. While early empirical research on identity (e.g. Phinney, 1990; Rumbaut, 1994) considered ethnic self-identification as a single linear variable, implying that feelings for the country of origin and for the host country are mutually exclusive, cross-cultural psychology studies (e.g Berry, 1980, 1984, 1997; Phinney, 1990) pointed to the possibility of coexistence between various ethnic and national identities, suggesting that the degree of identification to the majority culture should be treated as a separate concept from the degree of identification to the minority culture (e.g. Kvernmo and Heyerdahl, 1996; Pirie, 1996; Kinket and Verkuyten, 1997; Kolossov, 1999; Landale and Oropesa, 2002; Barrington, Herron and Silver, 2003; Bodenhorn and Ruebeck, 2003). Sen (1999) also highlighted that in the wider context of his multiple identities a person can reflect upon the importance he attributes to the adherence to a specific group. If minority identity and majority identity have a non-linear relationship, then the linear (oppositional) identity hypothesis may give misleading results. To allow for overlapping identifications, we will thus repeat our analysis for single as well as multiple identities.

We rely on a formal theoretical model to motivate our empirical specification, looking at the link between identity and language, and use these insights to deal with endogeneity problems in the estimation of the effect of identity on economic and political outcome variables. We offer numerous robustness checks for our hypothesized identity formation process (the first stage), as well as for the effects of identity (the second stage), allowing for single or multiple identities, looking at different sets of instruments, different subgroups and ethnic as well as religious identities.

3. Data

3.1. The TIES survey [4]

The dataset used in this paper is part of The Integration of the European Second Generation (TIES) project's survey of 10,000 respondents collected in 2007. We focus on two second-generation groups, Turkish and ex-Yugoslavian, and a 'native' group (individuals in the same age group whose parents were also born in the survey country) in the former guest-worker recruiting countries, Austria and Germany. These (second generation) groups constitute the two largest immigrant groups in Austria and Germany. The socioeconomic backgrounds of Turkish and ex-Yugoslavian labour migrants respectively are similar in both receiving countries, thus facilitating our cross-country approach. Similarities across countries are partly explained by the fact that migrants in different countries often originate from the same regions, or even the same villages. Furthermore, although there are also significant groups of refugees who fled the conflict between Kurds and Turks or as a result of the Yugoslav wars, most of them arrived in Austria/Germany later than the labour migrants, and their children are still young. Our sample thus consists of respondents who are almost exclusively children of labour migrants. This homogeneous composition of the sample is relatively unusual in the literature[5].

The respondents were between 18 and 35 years old. The second generation refers to those who were born in the receiving country, but at least one of their parents was not[6]. The same questionnaire was administered to a 'native' group, sampled as much as possible from the same neighbourhood where the second generation was sampled using the random route method (Kish, 1965). The 'native' group is thus not necessarily representative of the entire population of native youth in the city, however, given residential location,

---

[4]Information on sampling methodology is based on Groenewold and Lessard-Phillips (forthcoming).

[5]As also the receiving countries are similar in many respects, we prefer to exploit this homogeneity rather than focus on cross-country comparisons (also preferable given small sample sizes).

[6]Omitting those with one parent born in the survey country decreases the sample size from 798 to 701 but does not affect any of our findings. Information on parents and siblings is also derived from second generation respondents, respondents thus do not have to live with their parents for parental information to be available.

it is likely to be the group interacting with the second generation, affecting mutual relations, perceptions and attitudes.

One of the key challenges in obtaining representative samples of second generation immigrants is the identification of the sampling frame, i.e. the actual populations of second generation Turks and ex-Yugoslavs in the selected cities. For the countries examined here (Austria and Germany), existing administrative records could not be used to identify the second generation as they do not record the place of birth of the parents and also do not distinguish whether nationality is by birth or by naturalization. To overcome this constraint, municipal registers were used to compile a list of the forenames and surnames of all the inhabitants between the ages of 18 and 35 in the selected cities: Vienna, Linz, Berlin and Frankfurt (data was only collected in these four cities in Austria and Germany)[7]. Having constructed such full lists, names were then analyzed using onomastic software to derive ethnic origin and affiliation. Although there may be some omissions e.g. due to intermarriage, given the characteristics of migration history, we would not expect this to be of great importance[8]. The quality of such lists was examined by taking samples of names from the list and screening them by means of a short interview: errors seemed to be in the single-digit percentages. Although these limitations need to be kept in mind, we would not expect significant systematic bias in omitted groups.

The target groups were however difficult to contact and, once contacted, often refused cooperation for an interview (please see Table 1 below). Such low response rates immediately raise doubts about whether those who responded can represent those who did not in terms of personal characteristics and attitudes, in particular identity. The key problem that arises is that if non-respondents have, say, a stronger minority identity and are more likely

---

[7]While such municipal registers are probably more appropriate for the examination of the second generation than citizenship databases (containing only those with host country nationality, a non-random selection of second generation immigrants) or telephone directories (containing only those owning a landline, probably a similarly non-random sample), it should be remembered that they only list legal city residents. However, this is an acceptable shortcoming for our analysis, which looks at the determinants and effects of the identity of 'well-integrated' second generation youth.

[8]Also, both first and last names were considered. Historical migrations were taken into account (allowing for people with foreign last names who are not counted as migrants any more e.g. Huguenots and Poles in Germany) and variations in transcriptions were permitted (Humpert and Schneiderheinze, 2009).

to be unemployed/ lower educated, then the coefficient obtained from the restricted sample of respondents will be an underestimate of the true effect, i.e. will be biased downwards (vice versa for the 'opposite' correlation).

Table 1: Estimates of the reference population of study groups (the population), numbers of successfully interviewed study group members (the sample) and response rates by group, country and city

| | Austria | | Germany | |
	Vienna	*Linz*	*Berlin*	*Frankfurt*
Population				
Turkish	13,125	5,432	35,363	8,456
Ex-Yugoslavian	26,269	3,817	6,477	4,477
Comparison group	217,623	60,845	388,343	61,725
Sample				
Turkish	252	206	253	250
Ex-Yugoslavian	253	242	202	204
Comparison group	250	234	250	253
Response rate (%)				
Turkish	40.0%	70.0%	31.2%	24.8%
Ex-Yugoslavian	38.0%	38.0%	22.1%	22.9%
Comparison group	43.0%	42.0%	25.7%	24.3%

While in theory basic information on non-respondents could be derived from population registers and compared to the characteristics of respondents to examine the extent of selection bias, such information was not available for Austrian and German cities. Fortunately the questionnaire contained a question, to be completed by interviewers after each successful interview, on how difficult it was to get in contact with the respondent. The Continuum of Resistance model (Lin and Schaeffer, 1995; Stoop, 2005) asserts that difficult-to-reach respondents can be considered as a proxy for the unobserved non-respondents[9]. Estimates of the correlation for the difficult-to-reach group are thus used in section 5.3 to get a sense of which way our coefficients may be

---

[9]Interviewers were required to contact each selected respondent up to five times, recording which visit was successful and why earlier ones failed. The underlying assumption is that those who cooperated at the fifth visit are similar to those who maybe would have responded had they been visited a sixth time.

biased, to 'bound' the coefficient.

3.2. Sample descriptives

Table 2 provides the descriptive statistics by group, natives versus second generation immigrants. The mean age in the sample is around 26, with the second generation immigrants being slightly younger. The sample contains slightly more females than males as low response rates were particularly pronounced among young males. The natives are somewhat more educated and earn slightly more (on average between 1000 and 1500 Euros), and differences are more pronounced when looking at parental characteristics.

Table 2: Sample descriptives: summary statistics by group

	Variable	natives		second generation immigrants	
		Mean	Std. Dev.	Mean	Std. Dev.
personal characteristics	age	26.491	5.178	25.565	5.096*
	male	0.473	0.5	0.472	0.499
	education (1-4, in levels)	2.907	0.647	2.732	0.604*
	employed	0.847	0.361	0.767	0.423*
	income (1-9 categories)	3.342	1.344	3.002	1.094*
	survey country identity (1-6)	5.187	0.905	4.522	1.166*
	survey city identity (1-6)	4.941	1.041	4.556	1.096*
	European identity (1-6)	4.494	1.195	4.247	1.293*
	political orientation (1-5)	2.782	0.783	2.691	0.793*
parental char.s	father's age	54.969	7.228	54.296	7.122
	mother's age	52.228	6.835	50.951	6.537*
	father's education (1-4)	2.924	0.837	2.108	1.058*
	mother's education (1-4)	2.689	0.751	1.959	0.948*
	father employed when respondent 15 years old	0.947	0.224	0.94	0.238
	mother at home when respondent 15 years old	0.411	0.492	0.471	0.499*

Note: * denotes significant difference between natives and second generation immigrants at the 5% level

Table 3 compares Turkish and ex-Yugoslavian youth on a number of characteristics. We note the strikingly high percentage of second generation immigrants raised in German: 98 per cent and 95 per cent respectively, though

this could include schooling as well as language spoken with friends/ siblings, and does not necessarily refer to the language spoken with the parents[10]. While 94 per cent of the Turkish second generation respondents also report to have been raised in Turkish, only 33 per cent of the ex-Yugoslavs was raised in Serbian. (Note that while the ex-Yugoslav sample contains many ethnicities, focus in the following will be on the Serbian respondents as they constitute a relatively homogeneous group and make up by far the largest fraction.) Some of this disparity may thus be explained by a 'split' of the ex-Yugoslav sample among different languages.

Table 3: Sample descriptives: summary statistics by immigrant group

	Variable	Turkish second generation		ex-Yugoslavian second generation	
		Mean	*Std. Dev.*	*Mean*	*Std. Dev.*
identity	*Turkish/ Serbian identity*	4.783	1.342	2.764	1.941*
	Muslim/ Orthodox identity	4.509	1.631	2.383	1.883*
	raised in German	0.983	0.128	0.954	0.209*
lang. skills	*raised in Turkish/ Serbian*	0.941	0.236	0.335	0.472*
	can speak German	0.973	0.163	0.981	0.136
	can read German	0.896	0.306	0.964	0.185*
	can write German	0.849	0.359	0.925	0.263*
survey country citizenship		0.860	0.347	0.842	0.365

Note: * denotes significant difference between the Turkish and ex-Yugoslav groups at the 5% level

Although identification with the survey country is, as expected, higher among natives, the difference is surprisingly small, and almost negligible when comparing them to the Serbian group in terms of survey city or European identification. Of the two second generation groups Turkish respondents identify more strongly with their parents' country of origin, though this may again be driven partly by the 'split' of ex-Yugoslav identities among various country identifications. In fact most respondents in the ex-Yugoslav group feel strong associations with several groups, while identifying with the survey country/ city as well. In line with the earlier literature on overlapping

---

[10]The exact wording of the question was: 'In which language(s) were you raised?'.

identities, we find that a stronger Turkish or Serbian identity is in fact often associated with a stronger survey country identity, perhaps both driven by stronger 'political' (social?) views and awareness. This is a very important observation, which we will return to in the next section, since it suggests that relying on a simple linear 'either-or' model of identities can give strongly misleading results. Religious and ethnic identities do not necessarily 'move' together either - interactions between these identities will thus be considered as a robustness check. While the following analysis focuses on a pooled sample of the two immigrant groups, we have also examined them separately as a robustness check. The results were largely unchanged and are discussed in section 5.3.

Table A.1 in the Appendix reports summary statistics by language raised - significant differences here motivate our selection of control variables in the following sections.

Our outcome of interest in the first stage (identity formation) is the child's identity, constructed using the following survey question:

People can think of themselves as members of various groups in the wider society. The following questions are about how you think of yourself in this respect. I will read you a list of various groups in society. How strongly do you feel that you belong to these groups? To what extent do you feel...

- *[National]*

- *Turk/ [Ex-Yugoslav]*

- *[Inhabitant of city]*

- *European*

- *Muslim/ Orthodox*

- *[regional categories in country if relevant]*

- *[other minorities/ religious categories in country of parents' origin if relevant]*

Answer categories ranged from 'very strongly' through 'strongly', 'not strongly-not weakly', 'weakly', and 'very weakly' to 'not at all', and 'not applicable'. Note that the identity variable considered here is as reported by the children and as such is not necessarily identical to the one desired by the parents.

11

Our dependent variables in the second stage are education (measured on a 1-4 scale in levels: primary, lower secondary, upper secondary, tertiary), employment (a binary variable) and political orientation (measured on a 1-5 left-right scale)[11]. While there is a large literature on the education and employment outcomes of second generation immigrants, we have also included political orientation here as debates in the media are often centred on the perceived dangers of granting voting rights.

4. Theory

We are interested in the effect of identity on these economic and political outcomes, but since we are worried that identity might be endogenous we use an instrumental variables approach. The first issue this paper thus aims to investigate is identity formation, examining in particular how parental identity investments affect the child's identity. We then turn to a second stage, using these insights as a 'reduced form model of identity formation', supplying instruments for identity in regressions of economic and political outcomes on identity and a number of personal, parental and environmental characteristics. The models by Bisin et al. (2006) and Akerlof and Kranton (2000) serve as the theoretical framework behind our empirical methodology, highlighting the potential channels of identity formation as well as discussing why identity could affect outcomes.

4.1. The theory behind the first stage

We rely on the identity formation mechanism suggested by Bisin et al. (2006) where they model identity formation using an intergenerational model with parental investment in child identity. There is substantial psychological evidence for the importance of this channel (see e.g. Marks et al., 2007; Garcia Coll et al., 1996; Weiland and Coughlin, 1979; Erikson, 1968)[12] but they also allow for social interaction through peers and networks and an element of identity choice. This is in line with social identity theory, which emphasizes social interactions and self-esteem issues (see Tajfel and Turner, 1986).

---

[11]The exact wording of the question was: 'On a scale from politically left to right where would you put yourself?', with answer categories 'far left', 'left', 'middle', 'right' and 'far right'.

[12]For empirical work on intergenerational identity transmission see Rico and Jennings (2010); Duncan and Trejo (2009); Paryente (2008).

In the model suggested by Bisin et al. (2006) a population is composed of a majority ethnic trait (to which individuals can assimilate) and a minority trait. It is assumed that parents of the ethnic majority have children of the ethnic majority with no socialization effort, thus focus is on the decisions of the parents of the minority trait. Families are assumed to be composed of one parent and one child (both without gender) for simplicity. Children are born without defined preferences or cultural traits and are first exposed to their parent's trait. Cultural transmission inside the family to the parent's trait occurs with a probability increasing in costly socialization effort on the part of the parent. It is assumed that if a child from a minority family is not socialized in this way, he interacts with peers or role models in the neighbourhood in which he is raised and adopts the minority trait with a probability depending on the ethnic composition of the neighbourhood. In the third step the intensity of the child's ethnic identity is his personal choice. While our focus here is not on 'testing' the empirical implications of the model per se, we rely on it as a motivation for our reduced form estimation, looking in particular at the effects of parental investments. Note that the sequential nature of identity formation implied by this process is not necessary for the empirical tests, we may also think of the three channels as acting simultaneously.[13]

As the possibility of overlapping identities has been well-documented in the literature, we examine two specifications of the identity formation mechanism. First, we try to explain minority identity, constructed using the Turkish/ Serbian option of the survey question above[14]. Second, we examine

---

[13]Bisin et al. (2006) examined the model empirically using data from the UK Fourth National Survey of Ethnic Minorities on the parents' generation. While we build on the theoretical part of their paper, a different empirical approach is taken here (as discussed in detail in the following section). Although they have an intergenerational model in mind, due to data constraints they are restricted to examining parental identities. In contrast, we use data on the children's generation to examine the model's main outcome of interest, the child's identity. While they use the ethnicity of the partner as a measure of investments in the child's identity, we use language in which parents chose to raise their children as there is a large psychological and neurological literature on the link between language and identity (e.g. Joseph, 2004; Dong, 2009; Edwards, 2009; Feinberg, 2009) and this can be 'directly linked' to the children. We also explore the sensitivity of our results to a larger set of instruments.

[14]Unfortunately we cannot examine majority - Austrian/ German - identity separately as there is much less variation there and language raised is a weak instrument for majority identification.

a 'multiple identity' variable since we may believe that e.g. feeling strongly Turkish, while feeling strongly German at the same time is different from feeling strongly Turkish, but not German and we may want to allow for the effects of such interactions. Here we distinguish between four categories: a dominant minority identity, a dominant majority identity, two weak identities and two strong identities[15].

The second generation immigrants in our sample are 'split' roughly evenly among the four identity categories. In terms of raw means, those with a dominant majority identity are most educated and most likely to be employed among the four multiple identity categories, followed by those who have a strong majority as well as a strong minority identity (please see Tables A.2 and A.3 in the Appendix).

4.2. The theory behind the second stage

Our second stage is motivated by the recent theoretical work on identity in economics, in particular the papers by Akerlof and Kranton (2000, 2002, 2005, 2010) who see identity as a significant source of 'missing motivation'. While economists have traditionally seen identity as merely a 'label', which could at most act as a 'shifter' in the utility function, but is not a 'motivator', and therefore should not affect any choices, this question has received increased attention in recent years.

Akerlof and Kranton (2000) develop a framework based on the assumption that people have a view of who they are, a social category, and corresponding to who they are, their identity, they thus maintain an ideal for behaviour and lose utility insofar as they cannot live up to that ideal[16]. Akerlof and Kranton (2010) argue that 'In every social context, people have a notion of who they are, which is associated with beliefs about how they and others are supposed

---

[15]We created a binary variable out of each of the two categorical identity variables and looked at four categories (1,1), (1,0), (0,1) and (0,0) where e.g. (1,0) stands for feeling strongly German but not Turkish, etc. This approach is in line with the work of Berry (1980) or Constant, Gataullina and Zimmermann (2006).

[16]Although they make some reference to identity choice in their recent book (Akerlof and Kranton, 2010), the key shortcoming of their model is that the 'extended' utility function is simply taken as given and does not explicitly allow for identity choice, for the possibility that a person is able to reflect and to decide on who he wants to be. For explicit considerations of identity choice, see for example Sen (1985, 1999, 2006); Chen and Chen (2011); Fang and Loury (2005).

to behave' (p.4) - but how exactly would we expect identity to affect choices such as education, employment or political behaviour?

While current economic theories of education, for the most part, picture a student as a rational decision maker who weighs the economic costs and benefits of staying in school or choosing a particular occupation under considerable uncertainty, these decisions may be influenced by what individuals observe in their peer group, around them. As Akerlof and Kranton (2010) put it: 'How much schooling students get - what is called "the demand for education" - is largely determined by who they think they are and whether they should be in school' (p.15). Identity can thus influence 'real' choices and outcomes such as education or employment (how long to stay in school, how much effort to exert in school; which occupation to choose, how much effort to exert in job search or how high to set reservation wages) by affecting the reference group that the individual observes and compares himself to. This reference group could fulfil a dual function: (1) it provides a norm, a role model effect but (2) it also mediates the constraints of imperfect information as looking at the experiences of individuals 'close' to an individual is like running an experiment with better 'controls' and may therefore have better content in informing decisions[17]. A similar (perhaps more direct) link could be drawn between identity, reference groups and political behaviour.

Theoretically, the effect could go either way: a strong minority identity could encourage education, increase effort and lead to better employment outcomes if it is associated with aspirations for upward mobility in the host society (encouraging 'mobility investments')[18], however it could also act as an 'oppositional identity', incorporating a 'norm' for low educational effort, 'anti-school' values or opposition to school authorities. The empirical literature from the US found evidence of the latter among second generation immigrants, e.g. Waters (1994); Ogbu (1990); Fordham (1988); Portes and Zhou (1993) observed such 'oppositional' poses toward academic achievement among some American black teenagers ('acting white'). Such arguments could also be extended to employment (effort) choices. While we do not have a theoretical ex ante prior on which direction the effect on political orientation should go (in fact there may be a complex relationship

---

[17]Similar ideas on the role of aspirations were expressed by Appadurai (2004); Ray (2004); Mookherjee, Napel and Ray (2008); Bogliacino and Ortoleva (2010).

[18]This is similar to the minority group hypothesis originally put forward by van Heek (1956), see also e.g. Goldscheider (1967).

between views on immigration or integration policies and a general left-right orientation), we have included it as an outcome of interest since the paper was motivated by recent debates in Germany on the integration of (second generation) immigrants, many of which were not only concerned about the economic outcomes of the immigrants (or their effects on the host society, a question outside the scope of this paper), but also about political issues. The 2000 change in the German citizenship law and debates since have repeatedly addressed the question of 'Germanness', explicitly linking questions of identity to citizenship. The examination of this variable could thus constitute an interesting complement to the analysis of economic outcomes.

5. Estimation

We are interested in the effect of identity on outcomes such as education, employment and political orientation, however as we are worried about endogeneity, we begin with an empirical investigation of identity formation to provide insights for instrumental variables estimation.

Our first specification uses the language in which the respondent was raised as a single instrument since, as noted before, there is ample psychological and neurological evidence on the link between language raised and identity. Furthermore, language raised was a parental choice and is thus predetermined with respect to the respondent's current outcomes and identity. As we are concerned that having been raised in Turkish or Serbian could have a persistent direct effect on outcomes through worse German language skills (thus violating the exclusion restriction), we begin by restricting the sample to those whose German language skills (speaking, reading and writing) are 'very good' or 'excellent' (5-6 on a 1-6 scale). We then extend the sample to all second generation respondents to increase the sample size, but will explicitly control for German language skills. We believe that given that we are also controlling for a number of personal and parental characteristics as well as environmental influences and peer effects working through education (for details please see Table 4 below) our instrument should not affect the economic and political outcomes directly, but only through its effects on identity.

We start with this specification using only language raised as the single instrument as, given ample psychological evidence on the link between language and identity, we believe this to be the most plausible, but we also examine a number of parental characteristics (which we would not expect to

Table 4: Control variables

personal characteristics	*age, gender, number of siblings, whether has survey country citizenship, religion, German language skills*
parental/ family characteristics	*parents' education, whether the father was employed/ the mother was home when the respondent was 15 years old, siblings' education*
peer effects/ education	*whether attended kindergarten, whether any friends left education while in secondary school, proportion of children of immigrant origin in primary/ secondary school, type of secondary school (public/ private), whether has relatives in city of residence*
dummy variables	*group, city*

have a direct effect on the child's outcomes given our control variables) as instruments to assess robustness. Finally, we also look at results using this larger instrument set but excluding language raised as an instrument.

While our question would constitute a simple 2SLS problem with continuous variables, the situation here is complicated by the fact that our outcome is either an ordered categorical variable (education in levels, political orientation on a left-right scale) or binary (employment), identity is categorical (on a 1-6 scale for the minority identity model) and our instrument is also binary. We have opted for the use of a limited-information maximum likelihood (LIML) estimator, which allows us to specify the nature of these variables with a larger degree of flexibility[19]. While 2SLS necessarily assumes a linear first stage and ordered probit models do not allow the endogenous variable to be binary, this estimator allows for greater flexibility in estimation.

LIML gives consistent estimates that are equal to 2SLS estimates when an equation is exactly identified, whereas for overidentified equations, under standard assumptions[20], the LIML and 2SLS estimators are asymptotically equivalent and have the same asymptotic normal distribution. However, their

---

[19]We have implemented this in Stata using the 'cmp' command ('cmp' stands for Conditional Mixed Process); for further details on 'cmp' estimation please see Roodman (2011).

[20]fixed numbers of regressors and instruments, validity of the instruments, convergence of sample moments to population counterparts, and the ability to apply the central limit theorem

finite sample distributions differ; in particular LIML has better small sample properties than 2SLS: LIML tends to exhibit less bias than 2SLS and LIML confidence intervals typically have better coverage rates than 2SLS[21]. LIML thus has the advantage that while it has the same asymptotic distribution as 2SLS and gives asymptotic efficiency, it also provides a finite-sample bias reduction (Angrist and Pischke, 2009; Monte-Carlo study by Flores-Lagunes, 2007). We have estimated the model using both 2SLS and LIML - we gain increased confidence in our results since our findings from the two models are very similar. OLS results are reported for the effects of identity model for comparison.

5.1. First stage results

Table 5 below shows the first stage results. The first column reports the coefficient on language raised for a simple, single minority identity model (other control variables are as discussed in Table 4, for full results please see Table A.4 in the Appendix). Being raised in the minority language has a highly significant positive effect - in line with our theoretical predictions[22].

The second to fourth columns in Table 5 look at the determinants of identity when we explicitly allow for multiple feelings of belonging using a multinomial logit model. The dependent variable here is categorical: respondents can have a dominant majority identity, two weak identities or two strong identities (results are interpreted with respect to the base outcome: a dominant minority identity). Being raised in a minority language decreases the probability of identifying with the survey country relative to the country of origin of the parents; it also decreases the probability of having two weak identities compared to having a dominant minority identity - again in line with our theoretical predictions. Examining the last case of two strong identities, being raised in a minority language no longer has a significant effect - this is driven by the fact that this instrument is much worse at explaining majority rather than minority identification[23].

---

[21]LIML is also useful with many/ weak instruments as with many overidentifying restrictions 2SLS behaves poorly in finite samples.

[22]In fact looking at separate subsamples reveals that such effects are heterogeneous by age, with language raised having a larger effect on the identity of the younger than the older respondents, providing support for the plausibility of the model. These results are available from the author upon request.

[23]Turning to the roles of personal characteristics we observe a gender effect in line with

Table 5: First stage results

	minority identity	multiple ethnic identities			multiple ethnic-religious identities		
		dominant majority	*two weak*	*two strong*	*dominant religious*	*two weak*	*two strong*
raised in min.	0.8	-0.915	-1.54	0.03	-2.813	-1.952	-1.516
language	(0.127)**	(0.429)*	(0.430)**	(0.417)	(0.578)**	(0.465)**	(0.493)**
Number of obs.	798	798			796		
LR chi2	820.04	612.82			573.19		
Prob>chi2	0	0			0		
Log likelihood	-1461.78	-796.08			-701.83		

Note for all tables: * denotes significance at 5%, ** at 1%. Standard errors are reported in parentheses. Included covariates are as in Table 4. The 'minority identity' results have been estimated using the LIML estimator. The 'multiple ethnic identities' and 'multiple ethnic-religious identities' models have been estimated using a multinomial logit specification. The predicted probabilities for the 'multiple identities model' are: 0.252 (dominant minority identity), 0.217 (dominant majority), 0.131 (two weak), 0.237 (two strong). The predicted probabilities for the 'multiple ethnic-religious identities model' are 0.043 (dominant ethnic), 0.049 (dominant religious), 0.274 (two weak) and 0.399 (two strong). All results pertain to the full, pooled sample.

The last three columns in Table 5 report a similar multinomial logit model for combinations of ethnic and religious identities, as we may believe that e.g. feeling strongly Turkish and weakly Muslim is different from feeling strongly Turkish and strongly Muslim. The base category is a dominant ethnic identity, the other three categories are a dominant religious identity, two weak and two strong identities. Having been raised in the minority language makes a dominant ethnic identity more likely relative to all three

---

the earlier literature, with males being more likely to have a strong Turkish/ Serbian identity (please see Table A.4 in the Appendix). Having Austrian/ German citizenship significantly decreases minority identification. Being Orthodox significantly increases minority identification, while the Muslim dummy variable does not have a significant coefficient. Examining the role of parental characteristics, lower parental (in particular mothers') education increases the likelihood of having a minority identity. Turning to the influence of the external environment having attended kindergarten decreases minority identification, pointing to the role of early socialization. A larger proportion of children of immigrant origin in secondary school increases minority identification (as expected). Similar effects are observed when allowing for multiple feelings of belonging.

other options and is highly significant. The predicted values from this model will also be used in the second stage as a robustness check[24].

5.2. Second stage results

We now begin by looking at a restricted sample, including only those whose German language skills are 'very good' or 'excellent' to discount any direct effects of language raised on outcomes. Results are reported in Table 6 below. Our main finding is that once we control for endogeneity, identity does not have a significant effect on either of the economic or political outcomes: education, employment and political orientation[25].

Table 6: LIML and OLS results - 'very good' or 'excellent' German language skills

		education		employment		pol. orientation	
First stage	raised in min.			0.666			
results	language			(0.159)**			
		LIML	OLS	LIML	OLS	LIML	OLS
	minority	0.017	-0.045	0.065	-0.005	0.034	-0.029
Second stage	identity	(0.121)	(0.014)***	(0.189)	(0.012)	(0.120)	(0.038)
results	*Number of obs.*	543	543	536	536	428	428
	first stage F	15.26		14.6		7.27	

We then extend the sample to the entire pooled Turkish-Serbian second generation to increase sample size - Table 7 below reports these second stage results using LIML, as well as the OLS results for comparison. As findings are very similar to those above, having a minority identity again does not have significant effects on any of the economic or political outcome variables, we will rely on this full sample in further analysis.

Looking at the OLS results in Table 7 we can see the crucial importance of the endogeneity of identity: whereas our IV results show no significant effects of minority identity on outcomes, feeling strongly Turkish/ Serbian has a highly significant negative effect on education in the OLS model, even

---

[24]As the formation of ethnic and religious identities is not distinctively separable, we have also examined whether the results on multiple ethnic identities change by different religious identity categories. We obtained very similar results for both those with a strong and those with a weak religious identity

[25]Results are very similar when looking at only those with 'excellent' German language skills, but unfortunately there the sample size is even smaller.

	education		employment		political orientation	
	LIML	**OLS**	**LIML**	**OLS**	**LIML**	**OLS**
minority identity	-0.026	-0.042	-0.029	-0.011	-0.001	-0.001
	(0.095)	(0.012)**	(0.134)	(0.011)	(0.098)	(0.031)
Number of obs.	798	798	798	793	583	583
first stage F	33.20		32.32		18.85	

Table 7: LIML and OLS results

after controlling for a number of personal and parental characteristics and peer effects. This highlights the fact that an un-instrumented specification would overestimate the negative effect of a minority identity: once we control for endogeneity using an instrument, this effect is no longer significant.

While so far we have considered a restrictive specification using only one instrument, we also explore sensitivity to different sets of instruments. Table 8 compares the results using language raised as the single instrument with those using a larger set of instruments, including language raised as well as whether the respondents' parents were from a village or city and the length of the mother's stay in the host country (constructed as the mother's age now minus the mother's age at migration)[26]. Results are very much in line with those discussed above, enhancing the credibility of our IV estimates: crucially minority identity is still not significant for any of the outcomes in either specification.

As noted earlier, we also examine an instrument set that excludes language raised, using only the length of the mother's stay in the host country, whether the mother was from a village and a set of dummy variables for the province of origin of the mother as a robustness check. Results are again very similar (minority identity was still not significant for any of the outcomes) and are thus not reported, but are available from the author upon request.

Turning briefly to post-estimation checks, we gain increased confidence from the finding that all of our results have first stage F statistics, which very much exceed the rule-of-thumb of 10 (to avoid weak instruments, Stock, Wright and Yogo, 2002)[27]. The specification with several instruments also

---

[26]Unfortunately as this decreases the sample size a lot due to many missing values the LIML estimator could not be used, thus both specifications are reported relying on a simple linear 2SLS model.

[27]except for political orientation in the restricted sample and the 'four instruments'

Table 8: Second stage results using different instruments

	education		employment		political orientation	
	one instrument	*four instruments*	*one instrument*	*four instruments*	*one instrument*	*four instruments*
minority	0.011	0.004	-0.019	-0.035	0.155	-0.015
identity	(0.046)	(0.041)	(0.036)	(0.032)	(0.118)	(0.112)
Number of obs.	798	208	793	206	583	102
first stage F	33.20	14.26	32.32	13.24	18.85	9.61
overidentif. test p-value		0.59		0.13		0.81

passes the overidentification test.

5.3. Robustness checks

While two specifications have been explored for the first stage, looking at (1) the determinants of a minority identity and (2) the determinants of a multiple identity variable, the second stage results discussed above only used the first specification. If instead we explicitly allow for multiple identities, use the multinomial logit model in the first stage and correct for selectivity using predicted probabilities (Heckman, 1979), we get very similar results, with identity still not significant for either of the economic or political outcomes. This is also true when using the model, which allows for combinations of ethnic and religious identities (please see Table 9 below).

Table 9: Multiple identities - second stage results

	education	employment	political orientation
multiple (ethnic)	1.15	-0.097	-0.027
identity	(1.303)	(1.205)	(0.885)
Number of obs.	809	804	593
	education	**employment**	**political orientation**
multiple (ethnic-rel.)	0.184	2.387	1.047
identity	(1.304)	(1.577)	(1.531)
Number of obs.	809	804	593

---

specification, where the sample sizes are particularly small

We also examine the model separately by group (Turkish versus Serbian) since ethnicity may play a specific role in the formation of identity and its effects on outcomes (please see Table 10 below). The results are similar to the pooled effects, thus increasing our confidence. The significant negative effect of identity in the uninstrumented regression seems to be driven by the Turkish second generation group, where identity had a significant negative impact on education and employment outcomes.

Table 10: Robustness check - Turkish/ Serbian

		Turkish					
First stage	*raised in min.*	0.889					
results	*language*	(0.325)**					
		education		**employment**		**political orientation**	
		LIML	**OLS**	**LIML**	**OLS**	**LIML**	**OLS**
Second stage	*minority*	0.037	-0.072	-0.47	-0.036	-0.192	-0.007
results	*identity*	(0.202)	(0.023)**	(0.284)	(0.018)*	(0.222)	(0.055)
	Number of obs.	408		408		272	
	first stage F	8.13		6.29		5.15	
		Serbian					
First stage	*raised in min.*	0.705					
results	*language*	(0.150)**					
		education		**employment**		**political orientation**	
		LIML	**OLS**	**LIML**	**OLS**	**LIML**	**OLS**
Second stage	*minority*	-0.088	-0.026	0.16	0.014	-0.006	0.023
results	*identity*	(0.058)	(0.015)	(0.151)	(0.013)	(0.13)	(0.039)
	Number of obs.	390		390		311	
	first stage F	22.53		22.54		13.52	

Examining effects separately by gender (reported in Table A.6 in the Appendix) we obtain very similar results for men and women.

We find that difficult-to-reach respondents were more often male, less educated and more likely to be unemployed. We have thus examined both the identity formation mechanism and the effects of identity model separately for the difficult-to-reach group. Results confirm those for the easy-to-reach and are reported in Table 11 below: as before, minority identity does not have a significant effect on either of the economic or political outcomes and the link between identity and language raised is still highly significant and even larger in magnitude.

Table 11: Difficult-to-reach respondents

First stage	*raised in min.*		2.29	
results	*language*		(0.574)**	
		education	**employment**	**political orientation**
	minority	-0.185	0.287	-0.124
Second stage	*identity*	(0.198)	(0.303)	(0.199)
results	*Number of obs.*	110	110	110
	first stage F	15.73	15.22	21.41

Overall, we conclude that the insignificance of identity for economic and political outcome variables is a very robust result that holds across subsamples and specifications and using different sets of instruments.

6. Conclusion

This paper aims to provide insights into (1) the mechanisms driving the formation of identity and (2) the effects of identity on economic and political variables using empirical evidence from two second generation groups, Turks and ex-Yugoslavs in the former guest worker recruiting countries, Austria and Germany. Our main finding is the insignificance of identity for all three outcomes: education, employment and political orientation. We gain increased confidence in our results as they seem to be very robust to different instruments, subsamples and specifications (also allowing for ethnic and religious identity interactions) and also hold up once we deal with low response rates, examining difficult-to-reach respondents as a proxy for non-respondents.

Having been motivated by the debates in the media on (second generation) immigrants, we believe that the insignificance of identity for economic and political outcomes is in fact a very encouraging finding as it suggests that a strong ethnic or religious minority identity does not act as a constraint: feeling strongly Turkish/ Serbian does not hamper socioeconomic integration and does not affect political orientation.

Data sources

Austria

The data were made available by the Principal Investigator Barbara Herzog-Punzenberger.

Herzog-Punzenberger, Barbara (2010). Appendix: Stichprobendesign, Befragung und Evaluation. In: Herzog-Punzenberger, Barbara "40 Jahre und eine Generation später - die Kinder der angeworbenen Arbeitskräfte in Österreich sind erwachsen." Unpublished report to the Ministry of Science, Vienna. pp. 57-62.

Germany

The data were made available by Ms. Maren Wilmes, IMIS, University of Osnabrück.

Appendix

Table A.1: Sample descriptives: summary statistics by language raised

Variable	raised in minority language		not raised in minority language	
	Mean	*Std. Dev.*	*Mean*	*Std. Dev.*
personal characteristics				
age	25.426	5.110	25.820	5.064
male	0.489	0.500	0.439	0.497*
education	2.708	0.610	2.778	0.590*
employed	0.739	0.439	0.819	0.385*
income	2.926	1.036	3.132	1.175*
survey country citizenship	0.846	0.361	0.861	0.347
can speak German	0.973	0.161	0.983	0.129
can read German	0.910	0.287	0.965	0.185*
can write German	0.857	0.351	0.940	0.237*
survey country identity	4.374	1.209	4.795	1.029*
Turkish identity	4.875	1.274	3.185	1.530*
Serbian identity	4.425	1.451	1.868	1.541*
Orthodox identity	3.178	1.925	1.964	1.720*
Muslim identity	4.581	1.583	3.288	1.954*
political orientation	2.690	0.789	2.694	0.800
parental char.s				
father's age	54.320	7.383	54.252	6.624
mother's age	50.763	6.801	51.294	6.016
father's education	1.983	1.062	2.329	1.016*
mother's education	1.824	0.923	2.160	0.951*
father employed when respondent 15 years old	0.926	0.261	0.965	0.184*
mother at home when respondent 15 years old	0.520	0.500	0.381	0.486*
social networks				
attended kindergarten	0.723	0.448	0.779	0.415*
proportion of children of imm. origin in primary school	0.514	0.660	0.465	0.636
proportion of children of imm. origin in secondary school	0.552	0.690	0.416	0.608*
has friends who left education	0.382	0.486	0.235	0.425*
relatives in city of residence	0.896	0.305	0.865	0.342*

Note: * denotes significant difference between those raised in the minority language and those not raised in the minority language at the 5% level

Table A.2: Sample descriptives: multiple identity categories

	Percent
dom. minority identity	26.14
dom. majority identity	31.62
two weak identities	19.93
two strong identities	22.31

Table A.3: Sample descriptives: summary statistics by multiple identity categories

	dom. minority		dom. majority		two weak		two strong	
	Mean	*Std. Dev.*	*Mean*	*Std. Dev.*	*Mean*	*Std. Dev.*	*Mean*	*Std. Dev.*
education	2.573	0.624	2.856	0.581	2.707	0.618	2.776	0.542
employment	0.672	0.470	0.834	0.372	0.753	0.432	0.791	0.407
pol. orientation	2.682	0.781	2.733	0.843	2.675	0.702	2.609	0.811

Table A.4: First stage results using LIML

	education		employment		political orientation	
	Coef.	*Std. Err.*	*Coef.*	*Std. Err.*	*Coef.*	*Std. Err.*
raised in minority language	0.8	0.127**	0.791	0.128**	0.792	0.129**
age	-0.163	0.094	-0.16	0.095	-0.16	0.095
age2	0.003	0.002	0.003	0.002	0.003	0.002
male	0.179	0.083*	0.179	0.083*	0.179	0.083*
number of siblings	-0.046	0.044	-0.045	0.044	-0.045	0.044
has survey country citizenship	-0.465	0.124**	-0.46	0.124**	-0.461	0.124**
can speak German	-0.277	0.403	-0.276	0.403	-0.279	0.403
can write German	0.122	0.236	0.111	0.237	0.111	0.237
can read German	-0.175	0.317	-0.168	0.317	-0.167	0.317
Muslim	0.046	0.196	0.053	0.196	0.054	0.196
Orthodox	1.939	0.158**	1.95	0.158**	1.951	0.159**

(left margin label spanning rows: **personal characteristics**)

Continued on next page

		education		employment		political orientation	
		Coef.	*Std. Err.*	*Coef.*	*Std. Err.*	*Coef.*	*Std. Err.*
parental char.s	*father's education level*	-0.061	0.052	-0.062	0.052	-0.062	0.052
	mother's education level	-0.129	0.058*	-0.128	0.058*	-0.128	0.058*
	father employed when respondent 15 years old	-0.293	0.216	-0.29	0.216	-0.292	0.216
	mother at home when respondent 15 years old	0.067	0.095	0.062	0.095	0.062	0.095
	siblings' education levels	0.026	0.086	0.023	0.087	0.023	0.086
peer effects	*attended kindergarten*	-0.196	0.093*	-0.197	0.093*	-0.197	0.093*
	friends left educ.	-0.19	0.092*	-0.195	0.092*	-0.194	0.092*
	prop. of children of imm. origin in primary school	-0.072	0.091	-0.074	0.091	-0.073	0.091
	prop. of children of imm. origin in sec. school	0.204	0.091*	0.204	0.091*	0.204	0.091*
	type of secondary school (public/ private/ religious)	0.027	0.122	0.033	0.121	0.034	0.121
	has relatives in city of residence	0.126	0.156	0.127	0.156	0.127	0.156
Number of obs.		798		798		798	
LR chi2(52)		820.04		716.16		643.08	
Prob > chi2		0		0		0	
Log likelihood		-1461.777		-1303.594		-1602.262	

Note for all tables: * denotes significance at 5%, ** at 1%. Group and city dummy variables are not reported.

Table A.5: Second stage results using LIML

		education		employment		political orientation	
		Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
	minority identity	-0.026	0.095	-0.029	0.134	-0.001	0.098
personal characteristics	*age*	0.622	0.107**	-0.075	0.131	-0.026	0.11
	age2	-0.01	0.002**	0.001	0.002	0.001	0.002
	male	-0.078	0.096	0.767	0.123**	0.333	0.097**
	number of siblings	0.045	0.048	-0.004	0.057	0.069	0.05
	has survey country citizenship	-0.048	0.139	0.361	0.158*	-0.134	0.141
	can speak German	0.423	0.399	0.153	0.471	0.131	0.507
	can write German	0.152	0.262	0.793	0.281**	0.149	0.265
	can read German	0.62	0.346	0.006	0.379	-0.351	0.373
	Muslim	-0.101	0.199	-0.024	0.242	-0.196	0.185
	Orthodox	0.05	0.313	0.062	0.431	-0.283	0.311
parental char.s	*father's education level*	0.104	0.06	-0.104	0.071	0.027	0.058
	mother's education level	0.082	0.067	0.133	0.081	0.084	0.063
	father employed when respondent 15 years old	0.491	0.237*	0.503	0.259	-0.075	0.275
	mother at home when respondent 15 years old	-0.005	0.106	-0.237	0.126	-0.056	0.107
	siblings' education levels	0.114	0.097	0.039	0.115	0.102	0.097
	attended kindergarten	0.129	0.107	0.23	0.125	-0.018	0.107
peer effects	*has friends who left education while in secondary school*	-0.503	0.109**	0.155	0.126	-0.073	0.107

Continued on next page

Table A.5 – continued from previous page

		education		employment		political orientation	
		Coef.	*Std. Err.*	*Coef.*	*Std. Err.*	*Coef.*	*Std. Err.*
prop. of children of immigrant origin in primary school		-0.266	0.100**	-0.06	0.116	-0.17	0.106
prop. of children of immigrant origin in secondary school		-0.158	0.101	-0.257	0.119*	0.063	0.109
type of secondary school (public/ private/ religious)		0.187	0.129	0.339	0.176	0.412	0.121**
has relatives in city of residence		0.021	0.175	-0.123	0.222	-0.012	0.19
post-estimation tests							
First stage statistics	*adjusted R2*	0.603		0.602		0.652	
	first stage F	33.203		32.328		18.85	
	Prob>F	0.000		0.000		0.000	
Test of endogeneity, p-value	*chi2(1)*	0.223		0.792		0.157	

Table A.6: Robustness check - men/ women

		men					
First stage	*raised in min.*	1.148					
results	*language*	(0.237)**					
		education		employment		political orientation	
		LIML	**OLS**	**LIML**	**OLS**	**LIML**	**OLS**
Second stage	*minority*	0.006	-0.044	0.066	-0.017	0.061	-0.012
results	*identity*	(0.058)	(0.019)*	(0.038)	(0.012)	(0.150)	(0.052)
	Number of obs.	389		386		299	
	first stage F	23.48		23.3		16.03	
		women					
First stage	*raised in min.*	1.002					
results	*language*	(0.292)**					
		education		employment		political orientation	
		LIML	**OLS**	**LIML**	**OLS**	**LIML**	**OLS**
Second stage	*minority*	0.02	-0.04	-0.096	-0.009	0.28	0.01
results	*identity*	(0.073)	(0.017)*	(0.068)	(0.017)	(0.191)	(0.042)
	Number of obs.	409		407		284	
	first stage F	11.77		11.19		5.57	

References

Adsera, A., Chiswick, B.R., 2004. Are there gender and country of origin differences in immigrant labor market outcomes across European destinations? IZA Discussion Paper 1432.

Akerlof, G.A., 2007. The missing motivation in macroeconomics. American Economic Review 97 (1), 5-36.

Akerlof, G.A., Kranton, R.E., 2000. Economics and identity. Quarterly Journal of Economics 115 (3), 715-753.

Akerlof, G.A., Kranton, R.E., 2002. Identity and schooling: some lessons for the economics of education. Journal of Economic Literature 40 (4), 1167-1201.

Akerlof, G.A., Kranton, R.E., 2005. Identity and the economics of organizations. Journal of Economic Perspectives 19 (1), 9-32.

Akerlof, G.A., Kranton, R.E., 2010. Identity economics: How our identities shape our work, wages, and well-being. Princeton University Press, Princeton.

Angrist, J.D., Pischke, J.-S., 2009. Mostly Harmless Econometrics: An Empiricist's Companion. Princeton University Press, Princeton.

Appadurai, A., 2004. The capacity to aspire. In: Rao, V., Walton, M. (Eds.) Culture and public action. Stanford University Press, Stanford.

Barrington, L.W., Herron, E.T., Silver, B.D., 2003. The motherland is calling: Views of homeland among Russians in the near abroad. World Politics 55 (2), 290-313.

Battu, H., Zenou, Y., 2010. Oppositional identities and employment for ethnic minorities: Evidence from England. The Economic Journal 120 (542), F52-F71.

Berry, J.W., 1980. Acculturation as varieties of adaptation. In: Padilla, A.M. (Ed.), Acculturation: Theory, models and some new findings. Westview Press, Boulder, Colorado.

Berry, J.W., 1984. Multicultural policy in Canada: A social psychological analysis. Canadian Journal of Behavioural Science 16, 353-370.

Berry, J.W., 1997. Immigration, acculturation and adaptation. Applied Psychology: An International Review 46 (1), 5-34.

Bisin, A., Patacchini, E., Verdier, T., Zenou, Y., 2006. 'Bend it like Beckham': Identity, socialization and assimilation. CEPR Discussion Paper 5662.

Bodenhorn, H., Ruebeck, C.S., 2003. The economics of identity and the endogeneity of race. NBER Working Paper 9962.

Bogliacino, F., Ortoleva, P., 2010. Aspirations and growth: A model where the income of others acts as a reference point. Social Science Working Paper 1314R.

Casey, T. and Dustmann, C. 2010. Immigrants' identity, economic outcomes and the trasmission of identity across generations, The Economic Journal 120, F31-F51.

Chen, R. and Chen, Y. 2011. The Potential of Social Identity for Equilibrium Selection, American Economic Review, 101(6), 2562-89.

Clark, M.K., 2008. Identity among first and second generation African immigrants in the United States. African Identities 6 (2), 169-181.

Constant, A., Gataullina, L., Zimmermann, K.F., 2006. Gender, ethnic identity and work. IZA Discussion Paper 2420.

Crul, M., Schneider, J. 2009. Integration of Turkish second-generation men and women in Germany and the Netherlands: The impact of differences in academic tracking systems. In: Crul, M., Vermeulen, H., 2003. The second generation in Europe: Introduction. International Migration Review 37 (4), 965-986.

De Coulon, A., Wadsworth, J., 2008. On the relative gains to immigration: a comparison of the labour market position of Indians in the USA, the UK and India. Centre for Economic Performance Discussion Paper 851.

Dong, J., 2009. 'Isn't it enough to be a Chinese speaker': Language ideology and migrant identity construction in a public primary school in Beijing. Language and Communication 29 (2), 115-126.

Duncan, B., Trejo, S.J., 2009. Intermarriage and the intergenerational transmission of ethnic identity and human capital for Mexican Americans. Centre for Research and Analysis of Migration Discussion Paper 0902.

Dustmann, C., 1996. The social assimilation of immigrants. Journal of Population Economics 9 (1), 37-54.

Edwards, J., 2009. Language and Identity: An introduction. Cambridge University Press, Cambridge.

Erikson, E., 1968. Identity: Youth and Crisis. Norton, New York.

Fang, H., Loury, G.C., 2005. 'Dysfunctional identities' can be rational. American Economic Review 95 (2), 104-111.

Feinberg, T.E., 2009. From Axons to Identity: Neurological Explorations of the Nature of the Self. W.W. Norton & Company, New York and London.

Flores-Lagunes, A., 2007. Finite sample evidence of IV estimators under weak instruments. Journal of Applied Econometrics 22 (3), 677-694.

Fordham, S., 1988. Racelessness as a strategy in Black students' school success: Pragmatic strategy or pyrrhic victory? Harvard Educational Review, 58, 54-84.

Garcia Coll, C., Lawberty, G., Jenkins, R., McAdoo, H., Crnic, K., Wasik, B., et al. 1996. An integrative model for the study of developmental competencies in minority children. Child Development 67 (5), 1891-1914.

Georgiadis, A., Manning, A., 2009. Change and continuity among minority communities in Britain. Centre for Economic Performance Discussion Paper 903.

Goldscheider, C., 1967. Fertility of the Jews. Demography 4 (1), 196-209.

Groenewold, G., Lessard-Phillips, L., forthcoming. Research Methodology.

Heath, A., Cheung, S.Y., 2007. Minority ethnic disadvantage in the labour market: Britain. In: Heckman, J., Stixrud, J., Urzua, S., 2006. The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. Journal of Labor Economics 24(3), 411-482.

Heckman, J.J., 1979. Sample Selection Bias as a Specification Error. Econometrica 47, 153-161.

Humpert, A., Schneiderheinze, K., 2009. Sprachliche Analyse von Personennamen mit dem "Onomastik-Verfahren". Sozial- und Umfrageforschung. Duisburg, Germany.

Joseph, J.E., 2004. Language and Identity: National, Ethnic, Religious. Palgrave Macmillan, London.

Kinket, B., Verkuyten, M., 1997. Levels of ethnic self-identification and social context. Social Psychology Quarterly 60 (4), 338-354.

Kish, L., 1965. Survey Sampling. Wiley, New York.

Kolossov, V., 1999. Ethnic and political identities and territorialities in the post-Soviet space. GeoJournal 44 (4), 71-81.

Kvernmo, S., Heyerdahl, S., 1996. Ethnic identity in Aboriginal Sami adolescents: The impact of the family and the ethnic community context. Journal of Adolescence 19 (5), 453-463.

Landale, N.S., Oropesa, R.S., 2002. White, Black, or Puerto Rican? Racial self-identification among mainland and island Puerto Ricans. Social Forces 81 (1), 231-254.

Lewandowska, E., 2007. More Polish or more British? Identity of the second generation of Poles born in Great Britain. In: Zheng, G., Leung, K., Adair, J.G. (Eds.), Perspectives and Progress in Contemporary Cross-Cultural Psychology, Selected Papers from the Seventeenth International Congress of the International Association for Cross-Cultural Psychology.

Lin, I. F., Schaeffer, N.C., 1995. Using survey participants to estimate the impact of nonparticipation. Public Opinion Quarterly 59, 236-258.

MacFadden, M.A.J., 2004. Social insertion and identity construction of second generation Brazilian immigrants in Canada. Interfaces, Revista da ABECAN 4, 203-216.

Manning, A., Roy, S., 2007. Culture clash or culture club? The identity and attitudes of immigrants in Britain. Centre for Economic Performance Discussion Paper 790.

Marks, A.K., Szalachab, L.A., Lamarre, M., Boydd, M.J., Garcia Coll, C., 2007. Emerging ethnic identity and interethnic group social preferences in middle childhood: Findings from the Children of Immigrants Development in Context (CIDC) study. International Journal of Behavioral Development 31 (5), 501-513.

Mookherjee, D., Napel, S., Ray, D., 2008. Aspirations, segregation and occupational choice. BREAD Working Paper 187.

Nekby, L., Rodin, M., 2007. Acculturation identity and labor market outcomes. IZA Discussion Paper 2826.

Ogbu, J. U., 1990. Cultural model, identity and literacy. In: J. W. Stigler, R. A. Shweder, and G. Herdt (Eds.), Cultural psychology, essays on comparative human development, 520-541.

Paryente, B., 2008. Intergenerational values and identities transmission among German Christians, German Jews, Israeli Jews and Israeli Arabs. Paper presented at the ISPP 31st Annual Scientific Meeting, Sciences Po, Paris. Available at http://www.allacademic.com/meta/p254693_index.html.

Pendakur, K., Pendakur, R., 2005. Ethnic identity and the labour market. http://www.sfu.ca/~pendakur/pendakur_and_pendakur_ethnic_identity.pdf.

Phinney, J.S., 1990 Ethnic identity in adolescents and adults: Review of research. Psychological Bulletin 108 (3), 499-514.

Phinney, J.S., Horenczyk, G., Liebkind, K., Vedder, P., 2001. Ethnic identity, immigration, and well-being: an interactional perspective. Journal of Social Issues 57 (3), 493-510.

Pirie, P.S., 1996. National identity and politics in Southern and Eastern Ukraine. Europe-Asia Studies 48 (7), 1079-1104.

Portes, A., MacLeod, D., 1996. What shall I call myself? Hispanic identity formation in the second generation. Ethnic and Racial Studies 19 (3), 523-547.

Portes, A., and Zhou, M., 1993. The new second generation: segmented assimilation and its variants, Annals of the American Academy of Political and Social Science, 530(1), 74- 96.

Ray, D., 2004. Aspirations, poverty and economic change. BREAD Policy Paper 002.

Rico, G., Jennings, M.K., 2010. The intergenerational transmission of national identities: The case of Catalonia. Paper presented at the ISPP 33rd Annual Scientific Meeting, San Francisco. Available at http://www.allacademic.com/meta/p412581_index.html.

Roodman, D., 2011. Estimating fully observed recursive mixed-process models with cmp. Stata Journal 11 (2), 159-206.

Rumbaut, R.G., 1994. The crucible within: Ethnic identity, self-esteem, and segmented assimilation among children of immigrants. International Migration Review 28 (4), 748-794.

Schüller, S., 2011. Parental Ethnic Identity and Educational Attainment of Second-Generation Immigrants. IZA DP No. 6155.

Sen, A., 1985. Goals, commitment, and identity. Journal of Law, Economics, and Organization 1 (2), 341-355.

Sen, A., 1999. Reason before Identity. Oxford University Press, Oxford.

Sen, A., 2006. Identity and Violence: The Illusion of Destiny. Penguin Books, London.

Somerville, K., 2008. Transnational belonging among second generation youth: Identity in a globalized world. Journal of Social Sciences 10, 23-33.

Stock, J., Wright, J., Yogo, M., 2002. A survey of weak instruments and weak identification in Generalized Method of Moments. Journal of the American Statistical Association 20 (4), 518-529.

Stoop, I., 2005. The hunt for the last respondent: Non-response in sample surveys. SCP-report 2005/8.

Tajfel, H., Turner, J., 1986. The social identity theory of intergroup behavior. In: Worchel, S., Austin, W. (Eds.), Psychology of Intergroup Relations. Nelson-Hall, Chicago.

van Heek, F., 1956. Roman Catholicism and fertility in the Netherlands: Demographic aspects of minority status. Population Studies 10, 125-38.

Waters, M.C., 1994. Ethnic and racial identities of second-generation Black Immigrants in New York City. International Migration Review 28 (4), 795-820.

Weiland, A., Coughlin, R., 1979. Self-identification and preferences: a comparison of White and Mexican American first and third graders. Journal of Cross-Cultural Psychology 10 (3), 356-365.

Zephir, F., 2001. Trends in Ethnic Identification among Second-Generation Haitian Immigrants in New York City. Bergen and Garvey Publishers, Westport, Connecticut.

Zimmerman, L., Zimmerman, K.F., Constant, A., 2006. Ethnic self-identification of first generation immigrants. IZA Discussion Paper 2535.