Erschienen in: Journal of Public Economics. Elsevier. 2021, 200, 104448
https://doi.org/10.1016/j.jpubeco.2021.104448

More opportunity, more cooperation? The behavioral effects of birthright citizenship on immigrant youth ☆

Christina Felfe [a,b,d,i,*], Martin G. Kocher [c,d,e], Helmut Rainer [f,g,b], Judith Saurer [a], Thomas Siedler [h]

[a] *University of Würzburg, Germany*
[b] *CESifo, Germany*
[c] *University of Vienna, Austria*
[d] *IHS Vienna, Austria*
[e] *University of Gothenburg, Sweden*
[f] *University of Munich, Germany*
[g] *ifo Institute, Germany*
[h] *Universität Potsdam, Germany*
[i] *CEPR, United Kingdom*

ABSTRACT

Inequality of opportunity, particularly when overlaid with socioeconomic, ethnic, or cultural differences, may limit the scope of cooperation between individuals. A central question, then, is how to overcome such obstacles to cooperation. We study this question in the context of Germany, by asking whether the propensity of immigrant youth to cooperate with native peers was affected by a major integration reform: the introduction of birthright citizenship. Our unique setup exploits data from a large-scale lab-in-the-field experiment in a quasi-experimental evaluation framework. We find that the policy caused male, but not female, immigrants to significantly increase their cooperativeness toward natives. We show that the increase in out-group cooperation among immigrant boys is an outcome of more trust rather than a reflection of stronger other-regarding preferences towards natives. In exploring factors that may explain these behavioral effects, we present evidence that the policy also led to a near-closure of the educational achievement gap between young immigrant men and their native peers. Our results highlight that, through integration interventions, governments can modify prosocial behavior in a way that generates higher levels of efficiency in the interaction between social groups.

Keywords:
In-group favoritism
Out-group discrimination
Birthright citizenship
Lab-in-the-field experiment
Natural experiment

1. Introduction

Immigration has shaped, and continues to shape, many nations. This brings with it the challenge of integrating immigrants and their children into society. Among economists, one important take on integration—and on policies that promote it—is to emphasize convergence in the outcomes of immigrants and those of the host population in economic dimensions such as educational attainment and labor market participation (Algan et al., 2010; Sweetman and van Ours, 2014).

Another fundamental, but much less scrutinized, aspect of integration pertains to social interactions between immigrants and natives. Many such interactions, from everyday private exchanges to the provision of neighborhood amenities to working in teams, are not governed by enforceable contracts. Therefore, they almost always involve conflicts of interest and hold-up problems, and socially efficient outcomes will only be achieved if people are willing to cooperate.

However, socioeconomic, ethnic, and cultural differences between immigrants and natives can act as a barrier to cooperation. For example, inequality of opportunity increases the social distance between individuals, which has been shown to limit the

☆ We thank Sule Alan, Joshua Angrist, Manudeep Bhuller, Gordon Dahl, Natalia Danzer, Thomas Dohmen, David Figlio, Timo Hener, Albrecht Glitz, Stephen Jenkins, Edwin Leuven, Lucinda Platt, Carmit Segal, Steven Stillman, Uwe Sunde, Roberto Weber and numerous seminar and conference participants for useful comments. We wish to thank almost 20 research assistants and interns for their invaluable assistance with data collection and preparation. All errors remain our own.

* Corresponding author at: University of Würzburg, Germany.
E-mail addresses: christina.felfe@uni-wuerzburg.de (C. Felfe), martin.kocher@univie.ac.at (M.G. Kocher), rainer@econ.lmu.de (H. Rainer), judith.saurer@uni-wuerzburg.de (J. Saurer), thomas.siedler@uni-potsdam.de (T. Siedler).

scope of cooperation (Hoffman et al., 1996). It is also conceivable that individuals from disadvantaged groups adopt oppositional identities, which is said to involve "anti-social" behavior (Akerlof and Kranton, 2000). Thus, to think clearly about integration interventions, it is not enough to know about their impact in terms of educational or occupational outcomes; we should also be concerned about their potential in fostering cooperation between individuals of diverse backgrounds and perspectives.

To examine this issue, this paper zooms in on one fundamental mechanism for immigrant inclusion. Specifically, we study the effects of a major citizenship reform in Germany—the introduction of *birthright citizenship* on January 1, 2000—on the propensity of immigrant youth to cooperate with native peers. Our unique setup combines the advantages of experimental economics in studying in-group/out-group phenomena (Fershtman and Gneezy, 2001; Chen and Li, 2009) with the way in which labor economists have come to frame causal questions. In particular, we (i) conducted an incentivized lab-in-the-field experiment based on the investment (or "trust") game with a sample of more than 4,000 adolescents[1]; (ii) allowed participants to condition their strategies on the identity of their opponents; (iii) linked the experimental data with information from an extensive socioeconomic survey; and (iv) chose a sample design that allows us to connect the experiment with the citizenship reform using quasi-experimental identification strategies.

Birthright citizenship—the rule that all children born on a nation's soil obtain citizenship at birth—is subject to much controversy. For example, when Donald Trump became the first major U. S. presidential contender to endorse ending birthright citizenship, some saw it as an effective way of containing illegal immigration and birth tourism.[2] Yet others rallied to point out that birthright citizenship is one of the most powerful mechanisms of social inclusion (National Academies of Sciences, Engineering, and Medicine, 2015). This controversy, and similar debates in Europe, is surprisingly uninformed by reliable evidence from countries that have changed their regulations regarding birthright citizenship.

We exploit such a change and take a first step towards tracing its behavioral effects on young immigrants. Our lab-in-the-field experiment builds on the pioneering work of Fershtman and Gneezy (2001), who used the investment game (Berg et al., 1995) to study non-market interactions between real social groups. The advantage of using an incentivized experiment rather than a questionnaire is, as succinctly put by Fershtman and Gneezy (2001), that it captures people's behavior and not what people claim or believe to be their own behavior. We take the investment game as a vehicle to measure the extent to which individuals belonging to different groups are willing to cooperate, and the version we implement is based on the following idea. In a segmented migration society such as Germany, "being a native" or "being an immigrant" are amongst the core attributes of individuals (next to gender) that determine their social identities. Moreover, these attributes are ubiquitous and easy to perceive, and therefore, they are likely to feed into social interactions. Thus, using the strategy-vector method (Selten, 1967; Falk and Zehnder, 2013), we allowed participants to condition their decisions on the gender and migration background of their opponents.[3] Our main measure for intra- versus inter-group cooperation is the in-group/out-group investment gap that senders reveal in the first stage of the investment game. For natives it is the amount they send to other natives relative to the quantity they send to immigrants, while for immigrants it is the amount they send to other immigrants relative to the quantity they send to natives.

The main contribution of this paper is to examine the experimental data by exploiting natural variation induced by the following policy change. Until December 31, 1999, Germany granted birthright citizenship based on *jus sanguinis* (right of blood), i.e., only children born to German nationals received citizenship at birth. After January 1, 2000, the regime changed to a restricted version of *jus soli* (right of soil), i.e., every child born on German territory gained a conditional right to German citizenship. The conditionality attached to *jus soli* was that at least one parent had been a legal resident in Germany for eight years or more at the time of birth of the child. This setting provides us with a birth date eligibility cut-off, which serves as our source of identification. In particular, exploiting the quasi-random assignment of birthright citizenship around the cut-off, we analyze whether the policy caused discontinuities in immigrant children's propensity to cooperate with in-group and out-group members. To that end, we employ a local difference-in-differences (DID) model that compares immigrant children born shortly before and shortly after the cut-off date and draws upon native children as a control group. We conduct our analysis pooled by gender and separately for boys and girls because in-group/out-group phenomena of the type we study have been shown to exhibit gender differences (Fershtman and Gneezy, 2001). Moreover, the channels through which the introduction of *jus soli* might have affected immigrants' willingness to cooperate with natives—e.g., education, ethnic identity, or discrimination (more on this below)—naturally render themselves as candidates for gender-specific effects.

We obtain three main sets of results. Our core result concerns immigrants' behavior in the first stage of the investment game, and can be broken down into two parts: (i) the investments made by immigrant children born *pre-policy*, both boys and girls, reveals a marked gap between intra- and inter-group cooperation: on average, they transfer roughly 60% of their endowment to children with whom they share an immigrant identity, while their investments to native German children are 16% lower; (ii) the introduction of birthright citizenship significantly affected immigrant children's in-group/out-group behavior, but in a gender-specific way. In particular, we observe an in-group/out-group investment gap of 15% for immigrant boys born pre-policy, while for those born under *jus soli*, it is 11 percentage points lower. This implies that immigrant boys born under *jus soli* are almost equally inclined to invest toward immigrants and natives. This effect is entirely driven by an increase in out-group cooperation. For immigrant girls, the birth date cut-off does not matter at all: among those born pre-policy, investments to immigrants exceed investments to natives by 17%, and this difference persists for those born post-policy. Several robustness checks corroborate these results.

Second, in the investment game, there are two underlying motives for cooperating: the sender's beliefs about whether her choice to cooperate will be exploited by opponents (i.e., trust) and individual preferences such as other-regarding concerns (e.g., altruism) and risk aversion (see, e.g., Karlan, 2005; McEvily et al., 2012; Sapienza et al., 2013). As part of our design, we have elicited the expectations of senders regarding the back-transfer behavior of receivers, which we use as a proxy for trust. Moreover, given that we have employed the strategy method, we can use individuals' behavior as receivers as an indication of their other-regarding preferences. Exploiting these data, we find: (i) the introduction of birthright citizenship caused male, but not female, immigrants to

[1] As will be explained in more detail below, we have avoided selection in and attrition from the experiment by running it in 219 classes of 57 German schools during regular school hours in the final year of compulsory schooling. Throughout, we will use the terms children, youth and adolescents interchangeably to refer to the participants in our study.

[2] See, for example, a CNN article dated 18 August 2015, "Birthright citizenship: Can Donald Trump change the constitution?" (http://edition.cnn.com/2015/08/18/politics/birthright-citizenship-trump-constitution/index.html, accessed Oct 10, 2017)).

[3] For example, in the first stage of the investment game, participants had to decide how much of their initial endowment to send to a boy/girl with German-born/foreign-born parents.

significantly increase their trust towards natives; (ii) the policy did not affect immigrant children's other-regarding preferences towards natives. Moreover, we show that immigrant children do not hold wrong stereotypes of their native peers—i.e., they do not expect them to reciprocate less than they actually do—and the policy did not change this. These findings suggest that the increase in out-group cooperation among immigrant boys is an outcome of more trust rather than a reflection of stronger other-regarding preferences or less negative stereotypes towards natives.

Third, we explore factors that may explain these results. Since the increase in out-group cooperation among immigrant boys appears to be driven by an increase in out-group trust, and given that education has been argued to be the single best predictor of trust (Uslaner, 2008; Putnam, 2000), we first examine education as a potential mechanism for the reform's behavioral effects. This is a not-unlikely channel, since citizenship rights substantially improve, *inter alia*, immigrants' long-term economic perspective in the host country and may therefore act as a catalyst for human capital investments in immigrant families. We find, indeed, that he introduction of *jus soli* had strong educational effects for immigrant youth: it led to a near-closure of a substantial pre-existing educational achievement gap between them and their native peers. This finding is in line, and complements, our own previous research based on administrative school data (Felfe et al., 2020). Breaking the result down separately by gender, the reform's educational effect turns out to be entirely explained by male immigrants catching up educationally with their native peers.

A second plausible mechanism has to do with ethnic identity: it might be that citizenship causes immigrants to adopt the identity of "being German", and this in turn increases their willingness to cooperate with and trust people that share this identity. To explore this possibility, we draw upon our survey data to construct and examine an ethnosizer index that captures individuals' attachment to German society and culture in four domains: language, media usage, ethnic self-identification, and friendship network (Constant and Zimmermann, 2008). Interestingly, for both boys and girls, we find that birthright citizenship had a negligibly small impact on this ethnic identity measure, which speaks against the mechanism as an explanation for the reform effect.

Finally, it is possible that a discrimination channel is of relevance: natives might treat immigrants differentially based on their citizenship status, which may also affect immigrants' behavior towards them. To address this, we exploit our experimental design to examine whether natives treat their immigrant peers differentially based on their citizenship status. We find no evidence of such differential treatment in our experiment. However, as a caveat, our data does not allow us to examine whether naturalized and non-naturalized immigrants face differential experiences of discrimination by natives outside the laboratory.

Our results are important because they show that governments can modify and nurture prosocial behavior: the introduction of birthright citizenship in Germany brought about more cooperation between young immigrant men and their native peers and, consequently, higher levels of efficiency in the interaction between social groups. That said, our results also point to an important open puzzle: the positive reform effects we have uncovered are an entirely male phenomenon, i.e., the reform appears to have done little for the social integration of immigrant girls.

Our study relies on the combination of a lab-in-the-field experiment with a natural experiment and provides novel insights into the interface between immigration, citizenship and inter-group cooperation. It builds upon and connects a number of papers that span the fields of experimental economics, labor economics, and political science. Our research question is motivated by three findings from the extant literature on Germany's introduction of birthright citizenship. First, Avitabile et al. (2013) found that the policy caused immigrant families to increase their integration efforts. Second, Avitabile et al. (2014) provide evidence that the policy increased investments in child "quality" by immigrant parents. Third, in Felfe et al. (2020), a subset of us have shown that the policy led to a near-closure of the immigrant-native gap in a series of educational outcomes measured over children's early life cycle. These three findings have led us to formulate the hypothesis that the reform might have also had spill-over effects into the sphere of social interactions.

Our experimental design is an outgrowth of ideas developed by Fershtman and Gneezy (2001). In their experiments with Ashkenazi (Western) and Sephardi (Eastern) Jews in Israel, they find systematic discrimination towards men of Eastern origin. Inspired by this work, there has developed a small but active stream of literature in experimental economics on cooperation, trust, and discrimination between immigrants and natives. For instance, Guillen and Ji (2011) focus on domestic university students and their Asian international peers in Australia; Cox and Orman (2015) study first-generation immigrants and native-born Americans in the United States; and Cettolin and Suetens (2019) assess non-Western immigrants and native Dutch in the Netherlands. A study that is close in spirit to ours is Albrecht and Smerdon (2016). They exploit a refugee resettlement to a small rural town in Australia to study the effects of a migration shock on social capital. Combining trust data from a lab-in-the-field experiment with survey data from both treatment and similar control towns, they find that citizens in the treated town (i.e., who experienced the refugee resettlement) trust refugees relatively more than those in untreated towns.

The remainder of the paper is structured as follows. In Section 2, we focus on the lab-in-the-field experiment: we describe our setting, sample and experimental design, and provide basic results on intra- and inter-group cooperation among Germany's youth. Section 3 explains how we connect the lab-in-the-field experiment with the natural experiment of Germany's introduction of birthright citizenship. Section 4 contains the main results on the reform effect and its possible explanations. Section 5 concludes by offering some thoughts on policy implications and directions for future research. Four Appendices collect additional material.

2. The lab-in-the-field experiment

2.1. Motivation and key implementation challenge

The central idea of this study is to use experimental data on cooperation in a quasi-experimental evaluation framework. The specific question we ask is whether the propensity of immigrant youth to cooperate with native peers was affected by the introduction of birthright citizenship in Germany. A necessary condition for the implementation of our empirical approach is a large sample of children born in a narrow window around the reform's enactment date (January 1, 2000). To get at this group of children, and to ensure a large enough number of observations for immigrant youth in particular, we opted to run the experiment in schools. A close collaboration with two state ministries of education and school principals allowed us to collect data in 219 classes of 57 German schools during regular school hours. Given our empirical strategy, we restricted attention to a single school cohort, namely that composed of children mainly born in 1999 and 2000. When the experiment took place, all participants were in their final year of compulsory schooling, and thus 15 to 16 years of age.

2.2. Setting and subject pool

We first sought approval of and support for our study from educational authorities in Germany. The ministries of education of two

German federal states–Schleswig-Holstein (SH) and North Rhine-Westphalia (NRW)—approved our design and offered to support the project's implementation. Critically, this support included encouraging secondary schools in eight cities to participate in our project, which resulted in 57 school principals providing their agreement. The two federal states in which the study was run have independent education systems, which differ along one important dimension: in SH, the duration of compulsory schooling is nine years, while in NRW it lasts for ten years. In both federal states, a school year starts in August/September and ends in June/July. There were two waves of data collection. In the first wave, lasting from June 2 to July 15, 16, our target population were all 9th graders from 31 schools (spread over 122 classes) in six cities of SH.[4] In the second wave, lasting from October 19 to November 16, we targeted all 10th graders of 26 schools (spread over 97 classes) in two cities of NRW.[5]

The 57 participating schools belong to five school types: ten schools are secondary general schools ("Hauptschule"); eight are intermediate schools ("Realschule"); 29 are comprehensive schools without the final years of grammar school-type education ("Gesamtschule ohne gymnasiale Oberstufe"); eight are comprehensive schools with the final years of grammar school-type education ("Gesamtschule mit gymnasialer Oberstufe"); and two are grammar schools or high schools ("Gymnasium"). Two weeks prior to the study, school principals informed parents about the study and gave them an opt-out option, i.e., parents could proscribe their children's participation.[6] Moreover, immediately before the experiment started, all students present in class were informed by us that participation was voluntary. The experiment was run at the school class-level during two regular consecutive school hours.

On the days we conducted the study, a total of 4,634 students were present in the 219 classes. Parents made use of the opt-out option for 44 of them (less than 1%), while 154 students (3.5%) chose to opt out themselves. Thus, 4,436 students participated in the study. Of those, 133 participants did not fully complete the experimental task, while 226 did not provide the survey information necessary for our basic analysis (i.e., own gender and/or parental migration background). This leaves us with a baseline sample of 4,077 students.

The study consisted of two parts, the investment game (described in detail below) and an extensive socioeconomic survey. Each part lasted approximately one school period (45 min), and the order of the two parts was randomized on a daily basis in order to avoid any potential bias stemming from that sequence.[7] The study was conducted in regular classrooms and was done by paper and pen. To guarantee privacy, we installed mobile privacy screens between students.[8] We ensured anonymity by assigning a unique identity code to each participant.

2.3. Sample description

Our survey provides information, *inter alia*, about participants' date of birth, country of birth, citizenship, gender, school achievements, and family background. Two key family background variables are the birth places of both parents, which we use to categorize participants into three groups: (i) native children, whose parents are both German-born; (ii) immigrant children, whose parents are both foreign-born; and (iii) mixed-background children, who have one German-born and one foreign-born parent. Overall, according to our definitions, the sample comprises 2,201 native children (54%), 1,218 immigrant children (30%) and 658 mixed-background children (16%). Roughly 77% of all immigrant children in our sample are German-born (i.e., second-generation immigrants), while 23% are foreign-born (i.e., first-generation immigrants).

The empirical analysis contained in this paper mainly draws upon the experimental choices of immigrant and native children, while using that of mixed-background children only in passing. Thus, the following sample description is confined to the former two groups (for details, see Appendix Table A1). In Germany, the largest minority group of youth, by far, are immigrant children of Turkish origin. This is also evident in our sample. Specifically, 38% of immigrant children have parents who are Turkish-born, 14% have Middle-Eastern or African backgrounds, 12% have parents born in a post-Soviet country, 11% have parents from a Balkan country, 11% have Eastern European backgrounds, and 14% come from other countries. In terms of family characteristics, roughly one-third of immigrant children have parents with low educational attainment, three-fourths live in two-parent households, and more than one-half report a Muslim religious affiliation. When contrasted with same school-cohort data from the nationally representative German Microcensus, our sample of immigrant youth is comparable in terms of parental education, but it contains more children with a Turkish background (38% versus 31%; see Felfe et al., 2020, Table A2).

A comparison of native and immigrant children suggests several marked differences, of which we mention four. First, roughly one-third of immigrant children have parents with low educational attainment, while the corresponding share for native children is just under one-fourth.[9] Second, immigrant children are more likely than non-immigrant children to live in two-parent households (74% vs. 55%). Third, the majority of native children report a Christian (i.e., Catholic or Protestant) religious affiliation (67%), while the group of immigrants is predominantly made up of Muslim children (59%). Finally, 69% of immigrant children report that they speak a language other than German with their parents at home. This evidence reflects the pronounced cultural, social and economic gaps between native and immigrant children that are also observed in representative surveys.

2.4. The investment game: design and implementation

Our experiment is based on the standard investment game (Berg et al., 1995), which consists of two players, called the first-mover (sender) and the second-mover (receiver).[10] Each player is endowed with five euros at the beginning of the game. The first-mover decides on the amount to be sent to the second-mover ($x \in [0, 5]$) in steps of 50 €-cents. The transferred amount is then tripled by the experimenter. The second-mover can decide whether to send back any amount $y \in [0, 5 + 3x]$ to the first-mover. The final payoff for the first-mover is $5 - x + y$ and for the second-mover is $5 + 3x - y$. Under the assumption of selfish preferences, the only subgame-perfect equilibrium prescribes no investment and zero

[4] The cities are Flensburg, Kiel, Lübeck, Neumünster, Elmshorn, and Pinneberg, with population sizes ranging from 42,266 in Pinneberg to 246,306 in Kiel.

[5] The cities are Duisburg and Wuppertal, with population sizes of 491,231 and 350,046, respectively.

[6] Parents were, however, not informed about the objectives of the study.

[7] We pool the data of the two types of sessions in our main analysis. We discuss two robustness checks exploring how our experimental results depend on the order of events in the session (i.e., whether the survey or the experiment was conducted first).

[8] See Fig. B1 in Appendix B for a photo of a classroom setup.

[9] As Appendix Table A1 also shows, a relatively large proportion of immigrant children report that they do not know their parents' educational attainment.

[10] Public good games are also frequently used to study situations that require people to cooperate to achieve a goal that is considered beneficial to all. We have chosen the investment game because it has proven insightful in past research on non-market interactions between real social groups (Fershtman and Gneezy, 2001; Cettolin and Suetens, 2019). Moreover, and as we will next discuss, in our experiment we asked participants to decide both as first- and second-movers, which allows us to explore motives underlying cooperation.

returns. By contrast, "full" cooperation, where the first-mover invests his entire endowment, would maximize the players' joint payoff.

In our experiment, we employed the strategy method, i.e., each participant had to decide as first-mover and as second-mover. Moreover, and most importantly for the purpose of this paper, we allowed first-movers to condition their investment decisions on the gender and migration background of possible interaction partners. We implemented this by allowing first-movers to decide whether, and if so, how much, to transfer to each of six possible receiver types (indexed by k): a boy with German parents (S_1), a girl with German parents (S_2), a boy with foreign parents (S_3), a girl with foreign parents (S_4), a boy with foreign parents who possesses German citizenship (S_5), and a girl with foreign parents who possesses German citizenship (S_6).[11] In principle, this setup allows us to understand the extent to which cooperation is dependent on migration background as well as gender. However, the main task this paper sets itself is to examine intra- and inter-group cooperation between native and immigrant youth and how it is influenced by public policy. We will therefore largely abstract away from cooperation conditional on gender, apart from remarks, when deemed necessary.[12] Thus, we collapse the six choices $\{S_1, \ldots, S_6\}$ into two variables: a participant's average investment to natives (S_N) and his or her average investment to immigrants (S_I), defined as

$$S_N = \frac{1}{2}\sum_{k=1}^{2} S_k \quad \text{and} \quad S_I = \frac{1}{4}\sum_{k=3}^{6} S_k.^{13}$$

Throughout the paper, we refer to S_N as native children's in-group investments and as immigrant children's out-group investments, respectively. Likewise, we refer to S_I as native children's out-group investments and immigrant children's in-group investments, respectively. Our main measure for intra- versus inter-group cooperation is the in-group/out-group investment gap (IG) of senders with and without migration backgrounds. Formally, it is defined as

$$IG = \begin{cases} S_N - S_I & \text{for native children;} \\ S_I - S_N & \text{for immigrant children.} \end{cases}$$

Fig. 1 provides summary statistics on the investment choices of native and immigrant children by migration background of second-movers, both for the entire sample and separately by gender. Several interesting patterns emerge. First, for native adolescents, the evidence speaks against a strong pattern of unequal treatment of natives and immigrants. In the full sample [Panel (a)], natives' in-group investments exceed their out-group investments by a statistically significant 2.1% ($S_N = 2.90, S_I = 2.84$; paired t-test with $p < .01$). Looking at this result separately by gender [Panels (b) and (c)], we observe that there is no in-group/out-group variation in the investment choices of native girls ($S_N = 2.70, S_I = 2.72$). Native boys, by contrast, reveal moderate in-group favoritism: their in-group investments exceed their out-group investments by a statistically significant 4.8% ($S_N = 3.08, S_I = 2.94; p < .01$). Second, among immigrant children [Panel (a)], we detect a strong bias against natives that manifests itself in a statistically significant in-group/out-group investment gap of 13.4 percent ($S_I = 2.97, S_N = 2.62, p < .01$). The subgroup results by gender [Panels (b) and (c)] suggest that this gap is more pronounced for immigrant girls (16.1%; $p < .01$) than for immigrant boys (10.1%; $p < .01$). The main question of this paper, to which we turn in the next section, will be whether the introduction of birthright citizenship in Germany has caused discontinuities in immigrants' in-group/out-group behavior.

After participants had completed the first stage of the investment game, they were asked to indicate on their decision sheet the expected back transfer $E_k \in [0, 20]$ from the six possible interaction partners in steps of ten €-cents.[14]

At the final stage of the investment game, participants were asked to play the role of second-movers, and we employed the contingent response method to elicit their back transfers (returns). For example, on a first decision sheet, participants were asked to decide on their back transfers to a boy with German-born parents, contingent on the eleven possible investments of the boy as the first mover. Using the same strategy vector variant, we elicited back payments to the other five potential interaction partners. Amounts between and including 0 and $5 + 3x$ in steps of ten €-cents were allowed.

Before the experiment started, the instructions were distributed to all students in class and read out by an experimenter.[15] Students were informed that they would first play the investment game as the first-mover and thereafter as the second-mover. They were told that they could earn real money and that their payoffs would depend on their own choices and those of another, randomly assigned experiment participant from a different school.[16] The average payoff in the experiment was €7.26.[17] Participants received their payoffs no later than two weeks after the experiment took place (in envelopes with their unique identity codes, distributed by school secretaries or head teachers), which was known to them at the beginning of the experiment. All participants faced exactly the same decision tasks, instructions, and payoffs, and all procedures described here were common knowledge.

2.5. Discussion of experimental design

Several aspects of the experiment should be discussed. In light of our data requirements, the strategy method—i.e., asking participants to submit contingent decisions for native and immigrant opponents—was the only feasible way for collecting the experimental data. Specifically, through the strategy method, one obtains

[11] See Appendix B for the translated decision sheets.

[12] Appendix D provides the interested reader with analyses showing how the gender of interaction partners affects the results that we find. As becomes evident in this appendix, how cooperation depends on the gender of interaction partners is largely orthogonal to how it depends on migration background.

[13] Receiver types $k \in \{3, 4\}$ capture all immigrants (i.e., boys and girls with foreign parents), while receiver types $k \in \{5, 6\}$ capture only the subset of naturalized immigrants (i.e., those who possess German citizenship). The reason we have allowed for this distinction will become clear in Section 3, where we examine the effects of the German citizenship reform. For our main results, we have chosen not to drop any data, and hence, we compute S_I by averaging over their investments to receiver types $k \in \{3, 4, 5, 6\}$. That said, our results do not hinge on this specification, i.e., they remain qualitatively unchanged when we compute S_I by averaging over participants' investments to receiver types $k \in \{3, 4\}$, i.e., by letting $S_I = \frac{1}{2}(S_3 + S_4)$.

[14] We chose not to incentivize the elicitation of expectations for reasons of practicality.

[15] The translated instructions can be found in Appendix B. All sessions were conducted by one leading experimenter—in most cases, one of the authors—and one or two students assistants, previously trained by us. The experiment followed a strict protocol that was obeyed in every session.

[16] To be precise, participants were told that we would calculate final payoffs as follows: (i) we randomly match two participants from two different schools; (ii) we randomly assign the roles of first-mover and second-mover; (iii) we determine the true type k of both the first-mover and the second-mover based on survey information on own/other gender and whether parents are German-born or foreign-born; (iv) we implement the first-mover's decision for the true type of the second-mover; (v) we implement the second-mover's back transfer for the true type of the first-mover and his or her choice implemented in step (iv); and (vi) based on the pair of choices implemented in steps (iv) and (v), we calculate the participants' final payoffs. When we implemented this procedure to calculate participants' payoffs, we treated mixed-background children as children with foreign-born parents. Questions regarding the treatment of mixed-background children in the matching procedure were not raised by participants.

[17] The maximum payoff was €20. For participants whose payoffs were lower than €2, we paid out an unannounced consolation prize of €2.

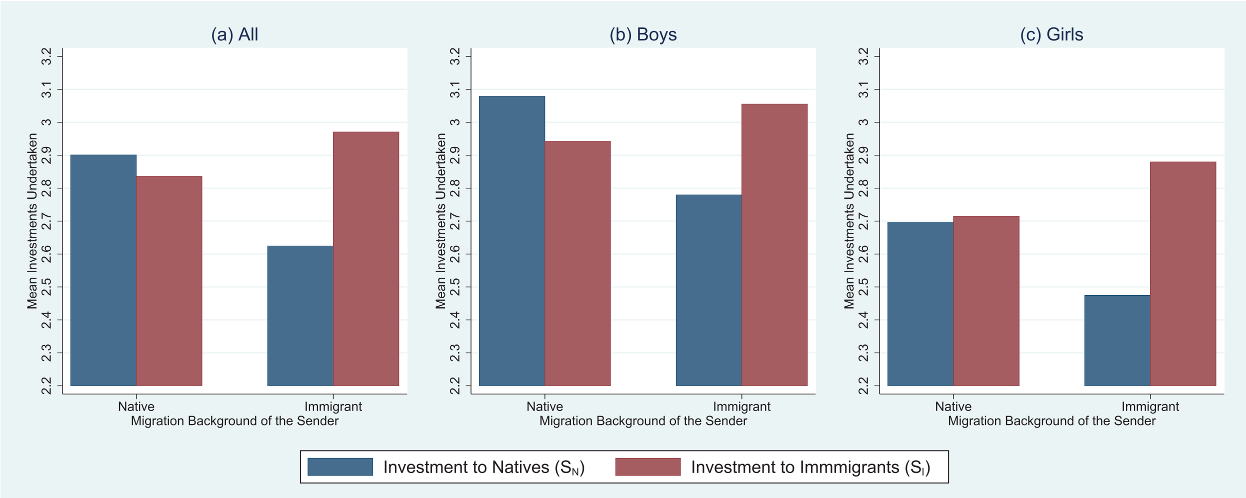

Fig. 1. First-Mover Investment Decisions of Native and Immigrant Children by Migration Background of Second-Movers.

for each participant a full set of strategies for all possible types of interaction partners, which keeps the required sample size within reasonable bounds. By contrast, through the direct-response method—i.e., randomly assigning to each participant either the role of sender or receiver and randomly matching him/her to one interaction partner of fixed type—one obtains only one decision per participant, which would have required us to draw a sample at least twice as large. Furthermore, the strategy method was the only feasible option in respect of data protection requirements, since it allowed for matching and payoff-calculation procedures entirely based on anonymous IDs. The direct-response method, instead, would have necessitated prior access to class lists containing personal data, which was incompatible with the data protection regulations set out by the cooperating school authorities.

Of course, having to submit strategies for both immigrant and native opponents may lead participants to think about decisions in a different way than had it been feasible to choose the direct-response method. For example, the potential role for experimenter demand biases (i.e., participants confirming or contradicting the experimenter's inferred hypothesis) may be larger. However, our aim is not to assess quantitative responses within our experiment, but to uncover qualitative results. That is, our interest does not lie in the *level* of in-group versus out-group investments *per se*, but in the behavioral effect of birthright citizenship on immigrant youth. To get at this effect empirically, we exploit natural, extra-experimental variation provided by a reform that led to a quasi-random assignment of birthright citizenship around a birthdate cut-off. Importantly, since our causing variable of interest—chil dren's birthdate—is orthogonal to the experimental design, participants could not form conjectures about our objectives. Thus, we are confident that experimenter demand does not generate or reverse a reform effect.

Finally, although our paper is not about a quantitative assessment of experimental responses, we make a brief note on the levels of elicited investments. A survey of the literature shows that the strategy method does not yield experimental results that differ systematically from those gathered through the direct-response method, especially if the stakes involved are high (Brandts and Charness, 2011). In our experiment, the monetary incentives were indeed substantial and should have made it costly to deviate from "true" preferred choices: data from the representative German Socio-Economic Panel suggest that the average payout in the experiment corresponds to more than 70% of the average amount of weekly pocket money given to adolescents with roughly the same characteristics as our participants.

Our estimation strategy in the next section will require us to narrow our full baseline sample, which contains the experimental choices of 4,077 participants. Thus, two self-contained Appendices have been included for readers interested in general (i.e., not reform-related) experimental results based on the full baseline sample. Appendix C demonstrates that general investment and back-transfer patterns in our experiment are comparable to what has been observed in previous, comparable experiments. Moreover, it provides detailed evidence on in-group/out-group investment patterns among Germany's immigrant and native youth, including a heterogeneity analysis for different immigrant groups.

3. The natural experiment

3.1. Institutional background: Jus Soli vs. Jus Sanguinis

The path to citizenship for immigrant children varies considerably across immigrant-receiving countries. In the United States, any person born on the nation's territory automatically gains U.S. citizenship, regardless of the nationality or immigration status of the person's parents. This rule, based on *jus soli* ("right of soil"), has been in place since the 19th century and is commonly referred to as birthright citizenship. By contrast, many countries in Europe have granted citizenship at birth based upon the principle of *jus sanguinis* ("right of blood"), meaning that citizenship is inherited through parents rather than determined by the place of birth. For children born to foreign nationals, this rule implies that citizenship can only be acquired through naturalization (i.e., upon application) later in life. Not surprisingly, in countries that have *jus soli*, virtually all native-born children of immigrants have the host-country nationality, while the lowest percentages of immigrant children with host-country nationality are found in countries that adhere to *jus sanguinis* (OECD, 2011).

The context of our study is Germany, a country that has recently witnessed a switch from *jus sanguinis* to *jus soli*. Throughout the 20th century, German citizenship could only be acquired by descent from a German mother and/or a German father. The legal status of immigrant children born to non-German citizens was either that of a temporary or a permanent resident. Although citizenship and permanent residency allow individuals to live in Germany indefinitely, they are very different statuses. Rights granted to permanent residents include the right to work in Germany and access to welfare benefits. However, permanent residents do not have the right to vote in general elections, are unable to apply for civil servant positions, cannot work in other EU countries, may lose their residency status if out of Germany for more than a year, and face the risk of deportation if they commit a crime. The two legal statuses, citizenship and residency, also have different implications for immigrants' labor market outcomes: compared with their non-naturalized peers, naturalized immigrants earn more (Chiswick, 1978; Steinhardt, 2012), have higher job-finding rates (Fougère and Safi, 2009; Gathmann and Keller, 2018) and experience steeper wage-tenure profiles (Bratsberg et al., 2002).

With the turn of the millennium, the principle of *jus sanguinis* was replaced by a restricted version of *jus soli*. In particular, every child born on German territory after December 31, 1999 gained a conditional right to German citizenship. The conditionality attached to *jus soli* was that at least one parent had been a legal resident in Germany for eight years or more at the time of birth of the child. For children born to parents who satisfied this residency criterion, German citizenship was *automatically* registered in the birth record without parents (i) having to apply for it and (ii) being able to disclaim it. Between January 1 and December 31, 2000, immigrant parents of children born between 1991 and 1999 were able to use a transition rule allowing them to *retrospectively* apply for their children's citizenship conditional on having legally resided in Germany for at least eight years. However, due to a lack of public information about this transition rule, only a small fraction of eligible families (roughly one-sixth) made use of it.[18] In the next section, where we outline our estimation strategy, we also present evidence that the reform had bite, i.e., due to the introduction of *jus soli*, a large portion of immigrant children were automatically endowed with German nationality at birth.

3.2. Empirical strategy: exploiting the natural experiment

We consider the introduction of *jus soli* in Germany on January 1, 2000, as an exogenous event that led to a quasi-random assignment of birthright citizenship among immigrant children. Since first-generation immigrant children (i.e., those born outside Germany) were unaffected by the reform, we exclude them from the analysis and only retain second-generation immigrant children (i.e., those born in Germany) for the estimation.

To isolate the effect the reform had on immigrant children's behavior, we compare the experimental decisions of second-generation immigrants born before and after January 1, 2000. In so doing, it is important to ensure that immigrant parents could not self-select into treatment. Since our source of identification is a birth date cut-off, the main concern is strategic fertility behavior. We address this issue in two ways. First, we restrict our sample to children born in the $\pm$4-month window around January 1, 2000. This ensures that our sample only comprises children who were conceived before July 1999, the month in which the German citizenship reform was ratified. In robustness checks, we further narrow the window around the reform cut-off date. Second, we implement a "donut" strategy that drops children born in the $\pm$2-week window around January 1, 2000. This avoids potential selection into treatment through birth-date-manipulation by parents.

Our analysis also needs to account for the fact that immigrant children born after the policy change are always younger than those born before it. Moreover, it is possible that the characteristics of parents change over the year (Buckles and Hungerman, 2013). To net out these potential sources of bias, we construct a second difference between pre-policy and post-policy native German children (who were unaffected by the reform cut-off date) and estimate the following local difference-in-differences (DID) model:[19]

$$\begin{aligned} \text{IG} = \gamma_0 + \gamma_1 \text{Immigrant} + \gamma_2 \text{Born Post - Reform} \\ + \gamma_3 (\text{Immigrant} \times \text{Born Post - Reform}) \\ + \theta \text{Birth Month} + \xi \text{Family} + \vartheta \text{Classroom} + \varepsilon, \end{aligned} \tag{1}$$

where *IG* refers to the in-group/out-group investment gap. The explanatory variable *Immigrant* is a binary variable indicating whether a child is a second-generation immigrant (=1) or a native (=0). The parameter γ_1 captures differences between immigrant and native children born prior to the policy change. *Born Post Reform* is a binary assignment variable indicating whether a child was born in the months just after January 1, 2000 (i.e., it is equal to one for children born between January and April 2000 and zero for children born between September and December 1999). The coefficient γ_2 measures general differences between children born before and after the citizenship reform that could cause changes in behavior even in the absence of a policy change. The coefficient of interest is γ_3, which multiplies the interaction *Immigrant*×*Born Post Reform* and thus identifies all immigrant children born after the policy change. We include a set of *Birth Month* dummies in all regressions. In extended specifications, we also include *Family* characteristics (i.e., maternal age, maternal education, family structure) and *Classroom* characteristics (i.e., class size, proportion of students with migration background, the gender ratio and five victimization measures[20]).

The coefficient of interest, γ_3, represents the reform's reduced form effect and can be interpreted as the intention-to-treat (ITT) effect of granting immigrant children citizenship at birth. This ITT effect is a conservative estimate of the impact of citizenship at birth, for two reasons: (i) our sample includes pre-policy children who may have qualified for citizenship at birth through *jus sanguinis*; and (ii) our sample includes not only immigrant children who were eligible for birthright citizenship when they were born but also those who were ineligible for it and were thus unaffected by the reform. The evidence on (i) and (ii) can be seen in Fig. 2, which draws upon two questions from our survey asking respondents about whether they hold German citizenship and since when they have held it. Panel (a) shows that, for all second-generation immigrant children in our sample, 35% percent of those born pre-policy have held German citizenship since birth, while the corresponding share for those born post-policy amounts to 78%. Put differently, the reform increased the share of second-generation immigrant children who have held German citizenship since birth by 43 percentage points. To gauge how conservative our results are, we will carry out a sensitivity check that exploits the fact that the reform had particularly strong bite for one immigrant subgroup.[21] Specifically, when the policy was implemented, a large portion of Turkish immigrants—an immigrant group which started to arrive in Germany via guest worker arrangements in the 1960s—ful-

[18] Children who acquired German citizenship via *jus soli* or the transition rule can hold two passports until the age of 23, when they have to opt for either German citizenship or that of their parents. When the cohort of immigrant children born in 1991 (who acquired German citizenship through the transition rule) had to choose between the two options in 2014, a large fraction opted in favor of their German citizenship (Worbs, 2014).

[19] Similar approaches have been used by Lalive and Zweimüller (2009), Dustmann and Schönberg (2012), Schönberg and Ludsteck (2014), Danzer and Lavy (2018) within the context of parental leave reforms.

[20] These victimization measures capture the proportions of students who report having been victims in the past year of physical abuse, verbal abuse, lies, theft or exclusion.

[21] Ideally, we would like to restrict our sample to eligible second-generation immigrant children, i.e., those whose parents fulfilled the residency criterion of eight years when they were born. However, data limitations prevent us from doing so. In particular, although our survey contains a question on parents' length of residence in Germany, a sizeable number of immigrant children report that they "don't know" their parents' residency duration. Consequently, restricting the sample on the available information on parents' length of residence in Germany would lead to a smaller and likely non-random subsample of immigrant children and would thus provide us with biased and imprecise estimates.

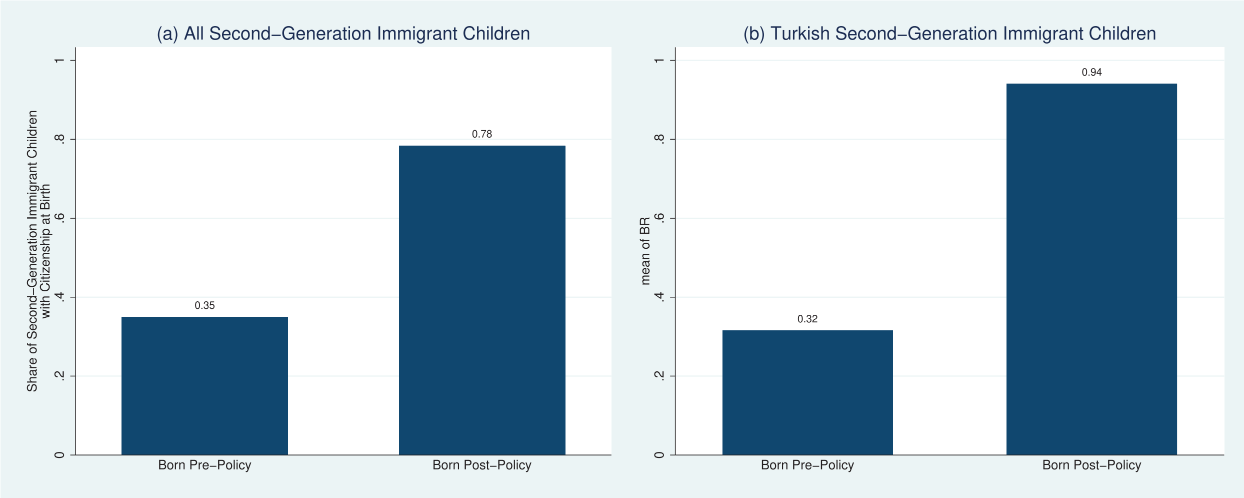

Fig. 2. Share of Second-Generation Immigrants with German Citizenship since Birth: A Comparison of Children Born Pre- and Post-Policy.

filled the residency criterion and became eligible for *jus soli*. This is confirmed in panel (b) of Fig. 2, which shows that 32% of all Turkish children born pre-policy have held the German citizenship since birth, while among those born post-policy this share increased by 62 percentage points to 94%.

Our full estimation sample comprises 920 native German children and 360 second-generation immigrant children. To verify whether treatment was balanced on observables, we present the mean values of key family and classroom characteristics for immigrant children born before and born after the reform in Appendix Table A2. The evidence shows that there are no systematic differences between children born before and after January 1, 2000. Among the 24 mean difference tests in both samples (see p-values in the last column of Table A2), only two mean differences are statistically significantly different from zero at the 5 percent level. One of these refer to mother's age, a difference that is to be expected given the reform's cut-off date. This supports the notion that the German citizenship reform was likely an "as-good-as-random" event with no systematic self-selection of particular types of immigrant families across the cut-off date.

4. Results

4.1. Cooperation in the investment game

In this section, we address our core question, namely, whether the propensity of immigrant youth to cooperate with native peers was affected by the introduction of *jus soli*.

Descriptive Evidence. We start by comparing the in-group/out-group investment behavior of second-generation immigrant children born pre- and post-policy. In Fig. 3, we regress our main outcome variable *IG* on a constant and the binary variable *Post-reform*, which indicates whether an immigrant child was born before (=0) or after (=1) 1 January 2000. The constant regression term captures the average in-group/out-group investment gap among immigrant children born pre-reform, while estimates of the coefficient on *Post-reform* shows how in-group/out-group investment gap of immigrant children born post-reform differs from those born pre-reform.

Let us first consider the results of the regression run for boys and girls together (labeled "All" on the x-axis). Among second-generation immigrant children born pre-policy, there is a marked gap between in-group and out-group cooperation: investments to immigrants (S_I = €3.00) exceed investments to natives (S_N = €2.59) by €0.41 or 16%. By contrast, among second-generation immigrant children born post-policy, this gap is €0.12 or roughly one-third lower (with S_I = €3.19 and S_N = €2.90), although this estimate is not statistically different from zero at 5% significance level. Behind this result is an interesting gender-specific pattern. Let us first consider the behavior of immigrant boys: Among pre-reform immigrant boys, investments to immigrants (S_I = €3.15) exceed investments to natives (S_N = €2.75) by €0.40 or 15%. By contrast, among post-reform immigrant boys, the in-group/out-group investment gap is a statistically significant 68%, or €0.27, lower. This means that immigrant boys born after the introduction of birthright citizenship are almost equally inclined to invest toward immigrants and natives. Decomposing the effect on the in-group/out-group investment into its components, we observe that the investments of immigrant boys to natives (respectively, immigrants) increases from a pre-policy mean of €2.75 (respectively, €3.14) to a post-policy mean of €3.24 (respectively, 3.37). Turning to the behavior of immigrant girls, it is interesting to observe that the birth date cut-off appears not to matter at all: among pre-reform immigrant girls, investments to immigrants exceed investments to natives by €0.42 or 17% (with S_I = €2.90 and S_N = €2.48), and this investment gap persists for post-reform immigrant females.

DID Estimates. The main concern with the results presented so far is that they may be confounded by age or season of birth effects. Thus, we now turn to our DID specification in Eq. 1. Before delving into regression results, it is important to consider the plausibility of our key identifying assumption. In particular, in analogue to the standard common trends assumption, our strategy requires that age-for-grade effects on children's behavior do not play out differently for immigrants and natives. Thus, in Fig. 4, we bin our main outcome variable by months of birth before and after the birthdate cut-off, and plot mean values for immigrant and native youth, respectively. We observe that, among children born pre-policy, the in-group/out-group investment gap is more pronounced among immigrants than among natives. In addition, and importantly for our identification strategy, the figure provides visual evidence of immigrant and native youth underlying common age-for-grade effects, and a treatment effect for immigrant boys (but not girls) that induces a sharp drop in the in-group/out-group investment gap at the birthdate cut-off.

In Table 1, we present estimates for Eq. 1, with and without the augmented set of control variables. In each Panel (A-C), the estimated coefficients in the first row ($\hat{\gamma}_1$) capture differences in in-group/out-group behavior between second-generation immigrant children and native children born prior to the policy change. The

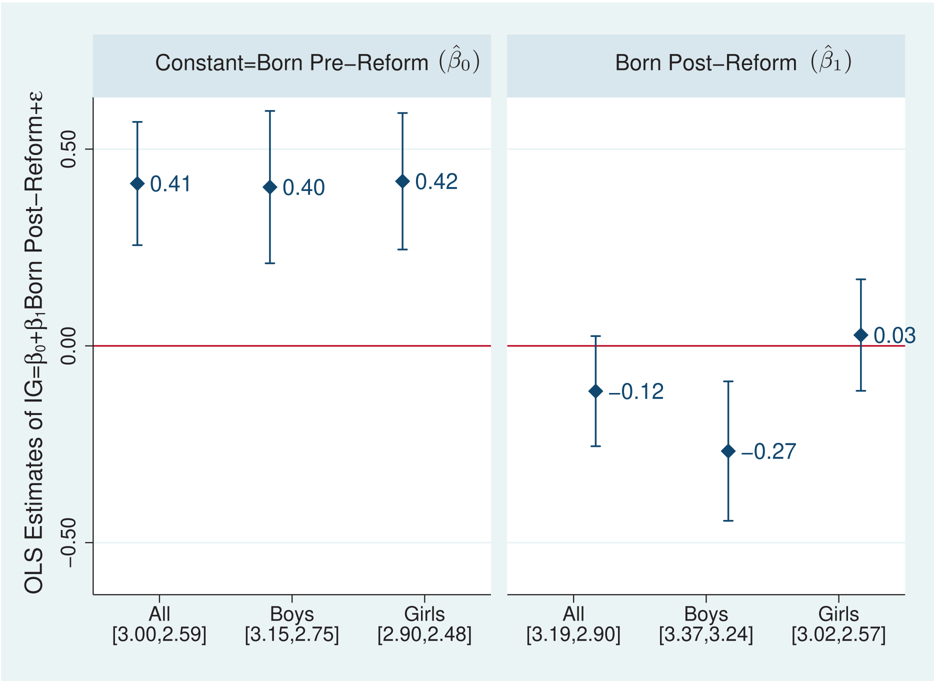

Fig. 3. In-Group/Out-Group Investment Gap Among Immigrants Born Around January 1, 2000. *Notes*: Sample comprises all immigrant children born between September 1999 and April 2000. ±2-week donut around the cut-off. Standard errors clustered by school type and school location. In square brackets, we report mean investments to immigrants (first entry) and mean investments to natives (second entry). Whiskers indicate the 95% confidence interval.

estimated coefficient of interest is $\hat{\gamma}_3$, which identifies the ITT effect of citizenship at birth on immigrant children's in-group versus out-group investments.

Let us first discuss the results of regressions run for boys and girls together (see Panel A). In Column (1), we only condition on gender and a full set of birth month fixed effects. For immigrant children born pre-policy, the mean of the in-group/out-group investment gap amounts to €0.41 (see Fig. 3). The estimate of 0.103 suggests that the introduction of birthright citizenship reduced this gap by approximately 25%, although this coefficient is not statistically significant at conventional levels. Columns (2) and (3) show the result to be robust to including controls for family background and classroom characteristics, respectively.

The remaining two panels of Table 1 break down the estimates by gender. Panel B presents the results for boys. Throughout all specifications, the reform effect turns out negative, is large in magnitude and is statistically significant at the 5% level. Consider our preferred specification in Column (3). The in-group/out-group investment gap among immigrant boys born pre-policy amounts €0.40, and the introduction of *jus soli* reduced this difference by €0.26, or 65%. In Panel C, where we repeat the analysis for immigrant girls, we find confirmation for the descriptive evidence presented above: the reduction in immigrants' in-group favoritism due to *jus soli* is an entirely male phenomenon. Irrespective of the specification, the reform effect for immigrant girls is small in magnitude—both in absolute terms and relative to estimates of $\hat{\gamma}_1$—and statistically indistinguishable from zero.

4.2. Robustness of the results

We now test the robustness of our main findings. All sensitivity checks, which are reported in Table 2, are conducted for our preferred specification [see Column (3) in Table 1].

Quasi-Eligible Sample. Due to the residency criterion attached to *jus soli*, the probability to be "treated" by the reform varied among second-generation immigrant children. The stand-out group were children with a Turkish background, for whom the reform had strong bite and therefore represents a particularly effective source of exogenous variation (see Fig. 2). Thus, compared to the results for the full estimation sample, we expect less conservative estimates if we restrict the treatment group to all second-generation immigrant children with a Turkish background and re-estimate our DID specification.

Column (1) of Table 2 presents the results, which confirm the intuition that the reform's effect is more pronounced among children with a Turkish background. In regressions run for boys and girls together [Panel A], the estimate of 0.254 suggests that the policy reduced the in-group/out-group investment gap among immigrants born-pre policy (€0.56) by 45%. Moreover, this reform effect is precisely estimated and differs from zero at the 5% significance level. The estimates by gender [Panel B and C] confirm that the reform effect is an entirely male phenomenon. For Turkish boys born pre-policy, we observe an in-group/out-group investment gap of €0.52, and the DID-estimate indicates a policy-induced reduction thereof of €0.57. This suggests that the reform induced Turkish boys to treat in- and out-groupers virtually equally. By contrast, the estimated coefficients for immigrant girls are small in magnitude and are statistically indistinguishable from zero.

Additional Robustness Checks. For our second robustness check, we recalculate our main dependent variable (IG) by letting $S_N = \frac{1}{2}(S_1 + S_2)$ and $S_I = \frac{1}{2}(S_3 + S_4)$ (see our discussion in footnote 13). Column (2) in Table 2 shows that the estimates based on this alternative outcome measure remain qualitatively unchanged compared to the benchmark results in Table 1.

In our analysis, standard errors are clustered by school type and school location, and there are a total of 18 clusters. Since reliable inference is a concern when there are few clusters (Donald and Lang, 2007; Cameron et al., 2008), our third robustness check tests whether the results also hold using wild cluster bootstrap *t*-procedures.[22] The estimates in Column (3) show that the p-values

[22] We estimated the wild cluster bootstrap standard errors using 1000 replications under H_1, as discussed in Cameron et al. (2008).

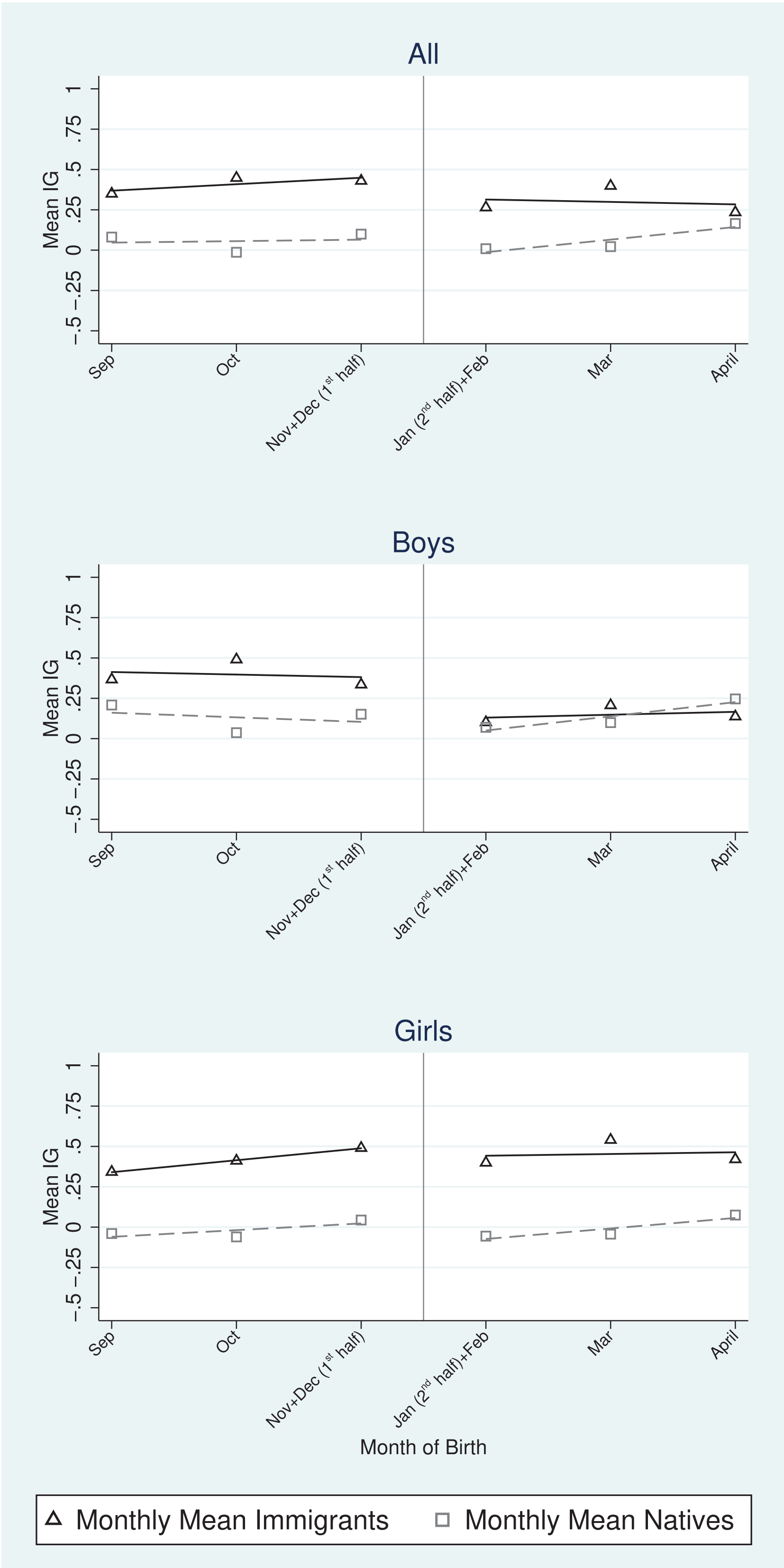

Fig. 4. In-Group/Out-Group Investment Gap (IG) Among Immigrant and Native Youth, Binned by Month of Birth Before and After January 1, 2000.

Table 1
Cooperation, DID Analysis.

Dependent Variable: In-Group/Out-Group Investment Gap (IG)			
	(1)	(2)	(3)
Panel A: All			
Immigrant ($\hat{\gamma}_1$)	0.360***	0.343***	0.337***
	(0.084)	(0.067)	(0.054)
Born post-reform*Immigrant ($\hat{\gamma}_3$)	0.103	0.099	0.106
	(0.078)	(0.073)	(0.071)
Observations	1,280	1,280	1,280
R-squared	0.038	0.048	0.056
Panel B: Boys			
Immigrant ($\hat{\gamma}_1$)	0.278**	0.314***	0.281**
	(0.109)	(0.100)	(0.112)
Born post-reform*Immigrant ($\hat{\gamma}_3$)	0.261**	0.284**	0.260**
	(0.106)	(0.106)	(0.094)
Observations	618	618	618
R-squared	0.017	0.049	0.062
Panel C: Girls			
Immigrant ($\hat{\gamma}_1$)	0.434***	0.427***	0.418***
	(0.083)	(0.085)	(0.066)
Born post-reform*Immigrant ($\hat{\gamma}_3$)	0.045	0.037	0.034
	(0.085)	(0.095)	(0.090)
Observations	662	662	662
R-squared	0.086	0.107	0.118
Month of Birth FE	Yes	Yes	Yes
Family Characteristics		Yes	Yes
Class Characteristics			Yes

Notes: OLS estimates of Eq. (1). Standard errors clustered by school type and school location and reported in parentheses; p-values reported in square brackets. All specifications in Panel A control for gender. *Family characteristics* include mother's age, dummy variables for mother's education (eight groups) and dummy variables for family structure (five groups). *Class characteristics* include class size, proportion of students with migration background, the gender ratio and five victimization measures (i.e., the proportion of students who report having been victims in the past year of physical abuse, verbal abuse, lies, theft or exclusion). *** (**) (*) indicates significance at the 1% (5%) (10%) level.

obtained from this bootstrap procedure confirm the levels of statistical significance reported in Table 1.

Recall that our study consisted of the investment game and a socioeconomic survey, and that we randomized the order of the two parts on a daily basis. In our fourth robustness check, we re-estimate our DiD model controlling for the order of two parts of our study. As can be seen in Column (4), this extended specification yields results virtually identical to those in Column (3) of Table 1. Thus, controlling for the order of the experiment and survey does not explain away the effect of birthright citizenship.[23]

The citizenship reform we study was ratified and announced in July 1999 but was already openly discussed in the German parliament during the previous month. Thus, our fifth robustness check provides estimations with a narrower $\pm$3-month-window around January 1, 2000. This additional restriction reduces the sample size by approximately 25%. Notwithstanding this, the results in Column (5) show that the coefficients remain largely unchanged compared to the benchmark estimates in Table 1, although they are less precisely estimated.

Our sixth robustness check involves placebo reform regressions. In particular, we shift the introduction of *jus soli* backward in time, assuming that it took effect on November 1, 1999. Moreover, we exclude all children born on or after January 1, 2000, from our placebo sample. The results in Column (6) show that the coefficients on the DiD interaction term are close to zero (or even positive) and statistically indistinguishable from zero. We conclude that the immigrant-native difference in intra- versus inter-group cooperation did not converge already among children born pre-policy.

4.3. Stereotypes, other-regarding preferences, and trust

The behavior that we observed in the first stage of the investment game may have an explanation based on wrong stereotypes: immigrant children may hold negative stereotypes of their native peers—i.e., they might expect them to reciprocate less than they actually do—and the introduction of birthright citizenship may have caused a drop in stereotyping. Our data reveals that this is not the case. Specifically, conditioning on immigrant children's investments to natives, and comparing their beliefs about back transfers with the actual back transfers they receive from natives, we find these expectations to be fairly accurate (see Appendix Fig. A1). Indeed, if anything, immigrant children, on average, slightly overestimate the amounts returned to them by their native peers. This result holds irrespective of whether we look at it separately by gender or by immigrant children born pre- or post-policy, respectively.

A potentially important motive for sending and returning money in the investment game are other-regarding preferences. Given that we have employed the strategy method, we can use individuals' behavior as receivers as an indication of their other-regarding preferences. To construct a measure of a participant's other-regarding preferences towards in-group versus out-group opponents, we proceed as follows. We let B_{km} denote an individual's back transfer to sender type $k = \{1, \ldots, 6\}$ who has send an amount $m = \{0, 0.5, 1 \ldots 4, 4.5, 5\}$, and calculate $P_k = \frac{1}{11}\sum B_{km}$, i.e., the receiver's back transfers to k averaged over the 11 possible investments from sender type k. Then, we collapse the six average back transfer $\{P_1, \ldots, P_6\}$ into the two variables, $P_N = \frac{1}{2}(P_1 + P_2)$ and $P_I = \frac{1}{4}(P3 + P_4 + P_5 + P_6)$, and then define in-group/out-group gap in other-regarding preferences as follows:

$$OG = \begin{cases} P_N - P_I & \text{for native children;} \\ P_I - P_N & \text{for immigrant children.} \end{cases}$$

In Column (1) of Table 3, we run Eq. (1) with OG as the dependent variable. The results can be summarized as follows. For immigrant boys born pre-policy, average back transfers to immigrants (P_I = €4.81) exceed average back transfers to natives (P_N = €4.51) by €0.28 or 6%, and the DID-estimate of 0.197 suggests that the introduction of birthright citizenship closed this small gap almost completely, although the estimate is far from reaching statistical significance. For immigrant girls born pre-policy, average back transfers to immigrants (P_I = €4.27) exceed average back transfers to natives (P_N = €4.21) by a mere €0.06 or 1%, and the statistically insignificant DID-estimate of 0.175 suggests, if anything, a quantitatively small increase in this gap. Taken together, the results suggest that the in-group/out-group gap in other-regarding preferences among immigrant youth is small, and that the introduction of birthright citizenship had no discernible effects on this gap.

Another important rationale for sending money in the investment game is trust, i.e., the sender's beliefs in the receiver's trustworthiness. Recall that we have asked participants to state expected back transfers E_k from the six possible interaction partners. This allows us to calculate a measure of a participant's belief about being exploited by in-group versus out-group opponents. To that end, we collapse the six expectations $\{E_1, \ldots, E_6\}$ into the two variables, $E_N = \frac{1}{2}(E_1 + E_2)$ and $E_I = \frac{1}{4}(E_3 + E_4 + E_5 + E_6)$, and then define the in-group/out-group gap in beliefs (BG) as follows:

$$BG = \begin{cases} E_N - E_I & \text{for native children;} \\ E_I - E_N & \text{for immigrant children.} \end{cases}$$

[23] Relatedly, asking individuals about citizenship in the survey before they participate in the investment game may be an example of priming citizenship. Reassuringly, when restricting our sample to children who played the investment game before taking the survey, our main conclusions remain the same. If anything, with this restricted sample, we find that the reduction in boys' in-group favoritism due to *jus soli* is larger than the estimated effect in our main specification in Table 1.

Table 2
Cooperation, DID Robustness Checks.

Dependent Variable: In-Group/Out-Group Investment Gap (IG)						
	Turkish subsample (1)	Alternative IG (2)	t-wild cluster (3)	Order of Experiment/Survey (4)	3-month window (5)	Placebo reform (6)
Panel A: All						
Immigrant ($\hat{\gamma}_1$)	0.497***	0.300***	0.337***	0.339***	0.330***	0.335**
	(0.126)	(0.059)	[0.000]	(0.055)	(0.085)	(0.121)
Born post-reform*Immigrant ($\hat{\gamma}_3$)	0.254**	0.112	0.106	0.112	0.069	0.048
	(0.089)	(0.082)	[0.263]	(0.073)	(0.095)	(0.158)
Observations	1,078	1,280	1,280	1,280	961	557
R-squared	0.06	0.05	0.06	0.06	0.07	0.112
Panel B: Boys						
Immigrant ($\hat{\gamma}_1$)	0.459***	0.262*	0.281**	0.288***	0.226	0.356**
	(0.157)	(0.130))	[0.040]	(0.114)	(0.146)	(0.161)
Born post-reform*Immigrant ($\hat{\gamma}_3$)	0.565***	0.282**	0.260**	0.272**	0.214	0.035
	(0.150)	(0.125)	[0.040]	(0.096)	(0.146)	(0.201))
Observations	529	618	618	618	461	265
R-squared	0.07	0.07	0.06	0.06	0.08	0.15
Panel C: Girls						
Immigrant ($\hat{\gamma}_1$)	0.586***	0.370***	0.418***	0.417***	0.454***	0.352**
	(0.156)	(0.057)	[0.000]	(0.066)	(0.085)	(0.149)
Born post-reform*Immigrant ($\hat{\gamma}_3$)	0.056	0.029	0.034	0.033	0.027	0.114
	(0.172)	(0.108)	[0.667]	(0.090)	(0.105)	(0.171)
Observations	549	662	662	662	500	292
R-squared	0.14	0.10	0.12	0.12	0.15	0.14
Month of Birth FE	Yes	Yes	Yes	Yes	Yes	Yes
Family Characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Class Characteristics	Yes	Yes	Yes	Yes	Yes	Yes

Notes: OLS estimates of Eq. (1). Standard errors clustered by school type and school location are reported in parentheses. For the specifications based on t-wild cluster bootstrap procedures, we report p-values in square brackets. All specifications in Panel A control for gender. *Family characteristics* include mother's age, dummy variables for mother's education (eight groups) and dummy variables for family structure (five groups). *Class characteristics* include class size, proportion of students with migration background, the gender ratio and five victimization measures (i.e., the proportion of students who report having been victims in the past year of physical abuse, verbal abuse, lies, theft or exclusion). *** (**) (*) indicates significance at the 1% (5%) (10%) level.

In Column (2) of Table 3 we estimate Eq. (1) with BG as the dependent variable. The main results are as follows. For immigrant boys born pre-policy, expected back transfers from immigrants (E_I = €5.56) exceed expected back transfers from natives (E_N = €4.89) by €0.67 or 14%, and the statistically significant DID-estimate of 1.103 indicates that the introduction of birthright citizenship not only closed this gap, but also changed its sign. Thus, the increase in out-group cooperation among immigrant boys (see Table 1) goes hand-in-hand with a substantial increase in trust towards natives. By contrast, for immigrant girls born pre-policy, expected back transfers from immigrants (E_I = €4.90) exceed expected back transfers from natives (E_N = €4.57) by €0.33 or 7%, and the statistically insignificant DID-estimate of 0.311 indicates, if anything, that the introduction of birthright citizenship increased this gap in beliefs.

4.4. Mechanisms and the null-effect for girls

In this section, we explore factors that may explain the behavioral effects of birthright citizenship. Note, however, that we cannot ascertain that these mechanisms are mutually exclusive, nor can we rule out that alternative mechanisms play a role.

Educational Effects. Since the increase in out-group cooperation among immigrant boys appears to be driven by an increase in out-group trust, and since education has been argued to be the single best predictor of trust (Uslaner, 2008; Putnam, 2000), we first examine education as a potential mechanism for the reform's behavioral effects. As mentioned at the outset, we consider this as an not-unlikely channel because endowing immigrant children with citizenship rights at birth can act as a catalyst for human capital investments in immigrant families. Indeed, in previous research based on administrative data, a subset of us (Felfe et al., 2020) has shown that the policy led to a near-closure of the immigrant-native gap in a series of educational outcomes measured over children's early life cycle (e.g., preschool enrolment, grade retention in primary school, tracking into secondary school).[24] We now extend these findings by exploiting information on school grades from our own survey.

In the German school system, grades vary discretely from 1 (excellent) to 6 (insufficient), with grades below 3 being considered achievements that exceed average requirements. We calculate a grade point average (GPA) based on participants' self-reported grades in the subjects "German" and "Mathematics". Then, we create the indicator *Above-Average GPA*, which equals one if a participant's average test score is better than 3, and zero otherwise.

In Panel A of Table 4, we run Eq. (1) with the indicator *Above-Average GPA* as dependent variable, both with and without the augmented set of controls. The key message one may extract from the table is that the citizenship reform substantially reduced the immigrant-native gap in school performance. For example, in our preferred specification [Column (3)], immigrant children born pre-policy are 9.9 percentage points less likely than their native peers (at an average of 41%) to achieve above-average grades. The statistically significant estimate of 8.9 percentage points suggests that the policy reduced this achievement gap by nearly 90%.

In Panels B and C of Table 4, in which the analysis is broken down by gender, it becomes evident that the leveling effect of the reform is entirely explained by immigrant boys catching up

[24] In Felfe et al. (2020), we argue that the educational effects due to *jus soli* are explained by stronger parental incentives to invest in children's human capital. Another possibility is that immigrant children with and without a German passport are treated differently by school teachers. In the two German states in which our study took place, parents do not provide information on their children's citizenship status when registering them at school. As such, neither principals nor teacher have *a priori* information on children's citizenship status. That being said, we cannot rule out that teachers acquire this information indirectly and act upon it.

Table 3
Other-Regarding Preferences and Beliefs, DID Analysis

Dependent Variable:	In-Group/Out-Group Gap in Other-Regarding Preferences	In-Group/Out-Group Gap in Beliefs
	(1)	(2)
Panel A: All		
Immigrant ($\hat{\gamma}_1$)	0.034	0.340
	(0.119)	(0.235)
Born post-reform*Immigrant ($\hat{\gamma}_3$)	0.015	0.274
	(0.163)	(0.289)
Observations	1,238	1,261
R-squared	0.017	0.033
Panel B: Boys		
Immigrant ($\hat{\gamma}_1$)	0.128	0.784**
	(0.135)	(0.282)
Born post-reform*Immigrant ($\hat{\gamma}_3$)	0.197	1.103**
	(0.233)	(0.456)
Observations	597	609
R-squared	0.037	0.065
Panel C: Girls		
Immigrant ($\hat{\gamma}_1$)	0.057	0.138
	(0.201)	(0.269)
Born post-reform*Immigrant ($\hat{\gamma}_3$)	0.175	0.311
	(0.200)	(0.447)
Observations	641	652
R-squared	0.029	0.064
Month of Birth FE	Yes	Yes
Family Characteristics	Yes	Yes
Class Characteristics	Yes	Yes

Notes: OLS estimates of Eq. (1). Standard errors clustered by school type and school location are reported in parentheses. For the specifications based on t-wild cluster bootstrap procedures, we report p-values in square brackets. All specifications in Panel A control for gender. *Family characteristics* include mother's age, dummy variables for mother's education (eight groups) and dummy variables for family structure (five groups). *Class characteristics* include class size, proportion of students with migration background, the gender ratio and five victimization measures (i.e., the proportion of students who report having been victims in the past year of physical abuse, verbal abuse, lies, theft or exclusion). *** (**) (*) indicates significance at the 1% (5%) (10%) level.

Table 4
Educational Achievement, DID Analysis

Dependent Variable: Above-Average GPA			
	(1)	(2)	(3)
Panel A: All			
Immigrant ($\hat{\gamma}_1$)	0.057*	0.058**	0.099**
	(0.029)	(0.024)	(0.040)
Born post-reform*Immigrant ($\hat{\gamma}_3$)	0.074**	0.076**	0.089**
	(0.031)	(0.035)	(0.035))
Observations	1,256	1,256	1,256
R-squared	0.009	0.025	0.045
Panel B: Boys			
Immigrant ($\hat{\gamma}_1$)	0.136***	0.148***	0.199***
	(0.044)	(0.038)	(0.056)
Born post-reform*Immigrant ($\hat{\gamma}_3$)	0.183***	0.184***	0.184***
	(0.050)	(0.058)	(0.060)
Observations	604	604	604
R-squared	0.014	0.035	0.052
Panel C: Girls			
Immigrant ($\hat{\gamma}_1$)	0.002	0.008	0.036
	(0.046)	(0.050)	(0.067)
Born post-reform*Immigrant ($\hat{\gamma}_3$)	0.016	0.014	0.014
	(0.052)	(0.054)	(0.048)
Observations	652	652	652
R-squared	0.005	0.029	0.067
Month of Birth FE	Yes	Yes	Yes
Family Characteristics		Yes	Yes
Class Characteristics			Yes

Notes: OLS estimates of Eq. (1). Standard errors clustered by school type and school location and reported in parentheses. All specifications in Panel A control for gender. *Family characteristics* include mother's age, dummy variables for mother's education (eight groups) and dummy variables for family structure (five groups). *Class characteristics* include class size, proportion of students with migration background, the gender ratio and five victimization measures (i.e., the proportion of students who report having been victims in the past year of physical abuse, verbal abuse, lies, theft or exclusion). *** (**) (*) indicates significance at the 1% (5%) (10%) level.

educationally with their native peers. For example, in Column (3) of Panel B, we see that immigrant boys born pre-policy are 19.9 percentage points less likely than native boys (at an average of 40%) to achieve above-average grades, but the statistically significant (at the 1% level) estimate of 18.4 percentage points implies an almost complete closure of this achievement gap. For immigrant girls (see Panel C), we obtain a different picture: the pre-reform immigrant-native achievement gap ($\hat{\gamma}_1$) among girls is lower than among boys, and the coefficients on the DID interaction are, though positive, statistically indistinguishable from zero. Overall, we conclude that the introduction of *jus soli* had the effect of bringing immigrant boys educationally on par with their native peers, while it had no such effect on immigrant girls.

Ethnic Identity. A second plausible mechanism is identity-related: it is possible that citizenship causes immigrants to adopt the ethnic identity of "being German", and this in turn might increase their willingness to cooperate with people that share this identity.

To operationalize the notion of ethnic identity in our context, we use information from our survey to construct an *ethnosizer index* (see, e.g., Constant and Zimmermann, 2008) that captures immigrants' attachment to German society and culture in four domains: (i) language; (ii) German media usage; (iii) ethnic self-identification; and (iv) friendship network. For (i), we create two dummy variables that take on the value 1 if a person communicates with her parents/friends only or mostly in German, and zero otherwise. For (ii), we create five binary indicators for whether an individual consumes or uses exclusively or mainly German media (books, internet, emails, television, movies). For (iii), we exploit the survey question "How much do you feel like a German (Very much, rather much, in some sense, not much, not at all)?", and create a dummy variable that equals one for participants who choose the answer categories "very much" or "rather much", and zero otherwise. Finally, for (iv), we create a dummy variable that equals 1 for individuals who report that at least half of their friendship network is comprised of Germans. Our ethnosizer index averages together the nine attachment measures in domains (i) to (iv).[25] The closer the ethnosizer index is to 1, the higher a person's attachment to German society and culture.

In Table 5, we re-run Eq. (1) but with the *ethnosizer index* as the dependent variable. The results can be summarized as follows. Non-reported summary statistics show that the ethnosizer index for native youth is close to unity, as should be expected (All: 0.923; Boys: 0.927; Girls: 0.919). Next, the highly significant estimate of 0.184 for $\hat{\gamma}_1$ in Column (3) of Panel A indicates that the ethnosizer index for immigrant youth born-pre policy is 20% lower than that of their native peers. Finally, the ITT-estimate of 0.004 for $\hat{\gamma}_3$ in Column (3) of Panel A shows that the introduction of *jus soli* had a null-effect on immigrants attachment to German society and culture, and this null-effect is fairly precisely estimated. Panels B and C, where we repeat the exercise separately by gender, confirm that this conclusion applies equally to boys and girls. Based

[25] As discussed by Kling et al. (2007), the aggregation improves statistical power to detect effects that go in the same direction for measures that capture young migrants' attachment to Germany society and culture.

Table 5
Social Identity, DID Analysis.

Dependent Variable: Ethnosizer Index	(1)	(2)	(3)
Panel A: All			
Immigrant ($\hat{\gamma}_1$)	0.227***	0.210***	0.184***
	(0.019)	(0.024)	(0.024)
Born post-reform*Immigrant ($\hat{\gamma}_3$)	0.002	0.003	0.004
	(0.028)	(0.030)	(0.028)
Observations	1,280	1,280	1,280
R-squared	0.326	0.342	0.363
Panel B: Boys			
Immigrant ($\hat{\gamma}_1$)	0.205***	0.192***	0.156***
	(0.032)	(0.036)	(0.033)
Born post-reform*Immigrant ($\hat{\gamma}_3$)	0.039	0.037	0.029
	(0.050)	(0.051)	(0.047)
Observations	618	618	618
R-squared	0.331	0.349	0.398
Panel C: Girls			
Immigrant ($\hat{\gamma}_1$)	0.240***	0.211***	0.194***
	(0.026)	(0.026)	(0.024)
Born post-reform*Immigrant ($\hat{\gamma}_3$)	0.022	0.019	0.014
	(0.037)	(0.039)	(0.039)
Observations	662	662	662
R-squared	0.336	0.370	0.383
Month of Birth FE	Yes	Yes	Yes
Family Characteristics		Yes	Yes
Class Characteristics			Yes

Notes: OLS estimates of Eq. (1). Standard errors clustered by school type and school location and reported in parentheses; p-values reported in square brackets. All specifications in Panel A control for gender. *Family characteristics* include mother's age, dummy variables for mother's education (eight groups) and dummy variables for family structure (five groups). *Class characteristics* include class size, proportion of students with migration background, the gender ratio and five victimization measures (i.e., the proportion of students who report having been victims in the past year of physical abuse, verbal abuse, lies, theft or exclusion). *** (**) (*) indicates significance at the 1% (5%) (10%) level.

on these findings, we consider ethnic identity as an unlikely channel for the reform effect on in-group versus out-group behavior.

Differential Treatment by Natives. The increase in out-group cooperation due to *jus soli* could be interpreted as a rational response by immigrant children if their native peers treat them differentially based on their citizenship status. Recall that in our design, opponent types $k \in \{3, 4\}$ refer to immigrants as a whole (i.e., boys and girls with foreign parents), while opponent types $k \in \{5, 6\}$ refer to the subset of naturalized immigrants (i.e., naturalized boys and girls with foreign parents). In Fig. 5, we analyze whether this distinction matters for the behavior of native children. In brief, the evidence suggests that native children do not systematically treat immigrants differently based on their citizenship status. For example, in Panel A, in which we illustrate the investment choices of native children as first-movers, we observe a small bias in favor of naturalized immigrants, but the investment gap to immigrants as a whole hardly exceeds 1%. In Panel B, where we look at the back transfers of native children as second-movers, the citizenship status of immigrant children appears not to matter at all. Finally, no gender-specific patterns can be observed.

Summary Thoughts on Mechanisms. For immigrant boys, the educational gap separating them from their native peers has a mirror image in terms of in-group/out-group cooperation and trust: Before the citizenship reform took effect, they were lagging behind their native peers educationally and strongly disfavored them, compared to other immigrants, in the investment game. The introduction of *jus soli*, in turn, saw a substantial increase in out-group cooperation among immigrant boys together with a near-closure of the achievement gap between them and their native peers. Besides these strong educational effects, we find no evidence that the reform caused young immigrants to adopt the ethnic identity of "being German", i.e., their attachment to German society and culture in domains such as language, media usage, self-identification, or friendship networks did not increase. Moreover, our experiment shows that native youth do not treat their immigrant peers differentially based on their citizenship status, suggesting that less discrimination by natives is an unlikely channel of reform effect on immigrant boys' out-group cooperation. That said, real-world experiences of discrimination by natives might be an important factor shaping immigrants' cooperation with natives even if their peers do not discriminate against them in the lab-in-the-field experiment. These findings provide us with a best-guess interpretation of the reform effect on young immigrant men: that the introduction of birthright citizenship caused them to increase their cooperativeness and trust towards natives because it led to a near-closure of the educational achievement gap between them and their native peers.[26] Although suggestive, several lines of thought underscore the potential relevance of this channel. First, individuals who have a high propensity to cooperate with and trust others tend to be optimistic about the future and feel they have control over their lives, and education is an important driver of this type of empowerment (e.g., Uslaner, 2008). Second, research in economics suggests that the lower the opportunities of disadvantaged groups to move upward socioeconomically, the less likely it is that members of these groups behave prosocially towards individuals from more advantaged groups (Akerlof and Kranton, 2000). In turn, the fact that birthright citizenship reduced educational gaps for immigrant boys suggests a potential relevance of this channel. Third, education positively affects cognitive skills (Carlsson et al., 2015), and cognitive skills are a prerequisite for rationally assessing if an object of trust merits trust (Lewis and Weigert, 1985; Hooghe et al., 2012). Finally, and as mentioned at the beginning of this section, other and not necessarily mutually exclusive mechanisms could also be at play. For example, Avitabile et al. (2013) find that immigrant parents are better integrated into society if their children are entitled to German citizenship at birth, and this may contribute to the behavioral and educational effects we have uncovered if boys benefit more from parental integration than girls do.

Immigrant Girls. For immigrant girls, our inquiry ultimately raises more questions than it answers. What explains their low out-group cooperation? Why has the introduction of *jus soli* worked for immigrant boys but not for immigrant girls? Although properly addressing these questions goes beyond the scope of this work, we believe that it opens up important new lines of research. For example, when examining parent-child dynamics in immigrant families, one frequently cited fact is that girls are often socialized to be "keepers of the culture" (Suárez-Orozco and Qin, 2006). Motivated by this observation, in theoretical and empirical follow-up research (Dahl et al., 2020), we explore the possibility that the intergenerational transmission of cultural identity makes an increase in economic opportunities an ineffective (or even contra-productive) policy lever for immigrant girls. Beyond that, it could also be that perceived discrimination by native peers is a salient factor in immigrant girls' out-group behavior. This seems important, not least because immigrant girls and immigrant boys often differ in external markers (e.g., wearing of headscarves) that may give rise to subtle and difficult-to-measure forms of discrimination. It would be interesting to also address this issue in future research.

[26] Of course, we cannot conclusively establish this as a key mechanism, as the reform we exploit does not render itself to a causal mediation analysis (Imai et al., 2013).

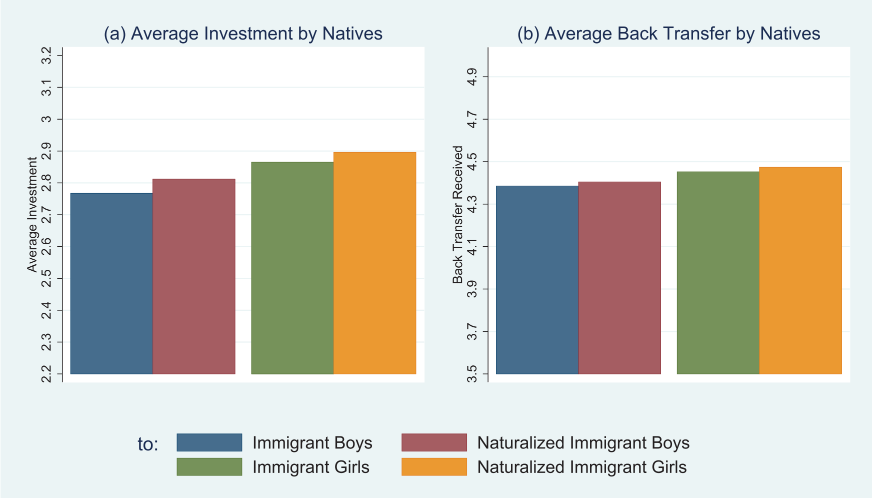

Fig. 5. First-Mover Investment Decisions and Second-Mover Back Transfers of Native Children by Gender and Citizenship Status of Immigrant Opponents.

5. Conclusions

Immigration has put many developed countries on a new demographic path. Immigrant children, in particular, make up a large and growing proportion of youth populations around the Western world. As a result, many scholars and policy makers argue that success in integrating immigrant children will be a crucial nation-building tool for years to come. Our starting point was the recognition that integration has several relevant dimensions. In particular, if we are to get a glimpse of the future face of Western societies, it is not just important to understand how today's immigrant children fare in the education system—we also need to know how children with and without migration backgrounds interact and whether integration policies can help overcome in-group/out-group phenomena and bring about cooperation between individuals with diverse backgrounds and perspectives. We have addressed this issue by combining a natural experiment—the introduction of birthright citizenship in Germany—with a lab-in-the-field experiment based on the investment game with more than 4,000 adolescents in their final year of compulsory schooling.

Descriptive evidence from the investment game shows that Germany's immigrant youth have a high propensity to cooperate with other immigrants but a low willingness to cooperate with their native peers. This suggests a need to reconsider some of the conventional wisdom about migrant integration. Discrimination *against* immigrants is an issue that figures prominently in many debates. Although we are not questioning the importance of this issue, it largely ignores the cleavages we have identified, i.e., that immigrants and natives may not be able to resolve social dilemmas because immigrants' cooperative decision-making is in-group bounded.

In connecting the experiment to the introduction of birthright citizenship, we have shown that these cleavages are not carved in stone. In particular, immigrant boys who, through the reform, received the same legal status as their native counterparts and as a result have caught up with them in terms of educational achievement appear to have extended their willingness to cooperate to their native peers. Thus, an important conclusion of our results is that governments can modify and nurture prosocial behavior, in our case resulting in more cooperation between immigrant males and native youth and, consequently, higher levels of efficiency in the interaction between social groups.

That said, the results also point to an important open challenge for policy makers: we have found that immigrant girls strongly discriminate in favor of immigrants and against natives in the investment game; yet, the positive reform effects we have uncovered—both in terms of out-group cooperation and education—are an entirely male phenomenon, i.e., the reform appears to have done little for the social integration of immigrant girls. This suggests that integration interventions are unlikely to offer "one fits all" solutions; those targeted at immigrant children may need to be gender-specific and take into account the different socialization pressures that immigrant girls and boys experience in the process of integration.

One issue we have not touched upon so far is whether our findings are generalizable to non-market interactions between immigrants and natives outside our laboratory setting and later in life as adults.[27] One interesting testing ground in this respect could be occupational settings. The cohorts born around January 1, 2000, will soon enter the labor market, where they will very likely encounter multicultural work environments. In such environments, immigrants and natives will need to cooperate, *inter alia*, as employees. With a suitably designed study, it would be feasible to analyze the scope for workplace cooperation between immigrants and natives and the long-run impact of birthright citizenship along this dimension. Thus, an important agenda for future research remains. Finally, our evidence comes from a single country of immigration, albeit the world's second-largest, and we prefer to avoid conjectures about external validity in this respect. Nonetheless, we believe that the results in this paper are useful for thinking about how widening the opportunities for disadvantaged groups may crowd-in social behavior that benefits society as a whole.

[27] The evidence about how trust develops over the life cycle is scarce and so far limited to two studies (Harbaugh et al., 2003; Sutter and Kocher, 2007). Both studies find a monotonic increase in transfers of trustors with age, which stabilizes during late adolescence.

References

Akerlof, G.A., Kranton, R.E., 2000. Economics and Identity. Quart. J. Econ. 115 (3), 715–753.

Albrecht, S., Smerdon, D., 2016. When Refugees Work: The Social Capital Effects of Resettlement on Host Communities. Unpublished Manuscript.

Algan, Y., Dustmann, C., Glitz, A., Manning, A., 2010. The Economic Situation of First and Second-generation Immigrants in France, Germany and the United Kingdom. Econ. J. 120 (542), F4–F30.

Avitabile, C., Clots-Figueras, I., Masella, P., 2013. The Effect of Birthright Citizenship on Parental Integration Outcomes. J. Law Econ. 56 (3), 777–810.

Avitabile, C., Clots-Figueras, I., Masella, P., 2014. Citizenship, Fertility, and Parental Investments. Am. Econ. J. Appl. Econ. 6 (4), 35–65.

Berg, J., Dickhaut, J., McCabe, K., 1995. Trust, Reciprocity, and Social History. Games Econ. Behav. 10 (1), 122–142.

Brandts, J., Charness, G., 2011. The Strategy Versus the Direct-Response Method: A First Survey of Experimental Comparisons. Exp. Econ. 14 (3), 375–398.

Bratsberg, B., Ragan Jr, J.F., Nasir, Z.M., 2002. The Effect of Naturalization on Wage Growth: A Panel Study of Young Male Immigrants. J. Lab. Econ. 20 (3), 568–597.

Buckles, K., Hungerman, D., 2013. Season of Birth and Later Outcomes: Old Question, new Answers. Rev. Econ. Stat. 95 (3), 711–724.

Cameron, A.C., Gelbach, J.B., Miller, D.L., 2008. Bootstrap-Based Improvements for Inference with Clustered Errors. Rev. Econ. Stat. 90 (3), 414–427.

Carlsson, M., Dahl, G.B., Öckert, B., Rooth, D.-O., 2015. The Effect of Schooling on Cognitive Skills. Rev. Econ. Stat. 97 (3), 533–547.

Cettolin, E., Suetens, S., 2019. Return on Trust is Lower for Immigrants. Econ. J. 129 (621), 1992–2009.

Chen, Y., Li, S.X., 2009. Group Identity and Social Preferences. Am. Econ. Rev. 99 (1), 431–457.

Chiswick, B.R., 1978. The Effect of Americanization on the Earnings of Foreign-Born Men. J. Polit. Econ., 897–921

Constant, A.F., Zimmermann, K.F., 2008. Measuring Ethnic Identity and its Impact on Economic Behavior. J. Eur. Econ. Assoc. 6 (2–3), 424–433.

Cox, J.C., Orman, W.H., 2015. Trust and Trustworthiness of Immigrants and Native-born Americans. J. Behav. Exp. Econ. 57, 1–8.

Dahl, G.B., Felfe, C., Frijters, P., Rainer, H., 2020. Caught between Cultures: Unintended Consequences of Improving Opportunity for Immigrant Girls. Working Paper 26674. National Bureau of Economic Research.

Danzer, N., Lavy, V., 2018. Parental Leave and Children's Schooling Outcomes: Quasi-Experimental Evidence from a Large Parental Leave Reform. Econ. J. 128 (608), 81–117.

Donald, S.G., Lang, K., 2007. Inference with Difference-in-Differences and Other Panel Data. Rev. Econ. Stat. 89 (2), 221–233.

Dustmann, C., Schönberg, U., 2012. Expansions in Maternity Leave Coverage and Children's Long-Term Outcomes. Am. Econ. J. Appl. Econ. 4 (3), 190–224.

Falk, A., Zehnder, C., 2013. A City-Wide Experiment on Trust Discrimination. J. Public Econ. 100, 15–27.

Felfe, C., Rainer, H., Saurer, J., 2020. Why Birthright Citizenship Matters for Immigrant Youth: Short- and Long-Run Impacts of Educational Integration. J. Lab. Econ. 38 (1), 143–182.

Fershtman, C., Gneezy, U., 2001. Discrimination in a Segmented Society: An Experimental Approach. Quart. J. Econ. 116 (1), 351–377.

Fougère, D., Safi, M., 2009. The Effects of Naturalization on Immigrants' Employment Probability (France, 1968–1999). Int. J. Manpower 30 (1–2), 83–96.

Gathmann, C., Keller, N., 2018. Access to Citizenship and the Economic Assimilation of Immigrants. Econ. J. 128 (616), 3141–3181.

Guillen, P., Ji, D., 2011. Trust, Discrimination and Acculturation: Experimental Evidence on Asian International and Australian Domestic University Students. J. Socio-Econ. 40 (5), 594–608.

Harbaugh, W.T., Krause, K., Vesterlund, L., 2003. Trust in Children. In: Ostrom, E., Walker, J. (Eds.), Trust and Reciprocity: Interdisciplinary Lessons from Experimental Research. Russell Sage Foundation, pp. 302–322.

Hoffman, E., McCabe, K., Smith, V.L., 1996. Social Distance and Other-Regarding Behavior in Dictator Games. Am. Econ. Rev. 86 (3), 653–660.

Hooghe, M., Marien, S., de Vroome, T., 2012. The Cognitive Basis of Trust. The Relation between Education, Cognitive Ability, and Generalized and Political Trust. Intelligence 40 (6), 604–613.

Imai, K., Tingley, D., Yamamoto, T., 2013. Experimental Designs for Identifying Causal Mechanisms. J. Roy. Stat. Soc. Ser. A (Stat. Soc.) 176 (1), 5–51.

Karlan, D.S., 2005. Using Experimental Economics to Measure Social Capital and Predict Financial Decisions. Am. Econ. Rev. 95 (5), 1688–1699.

Kling, J.R., Liebman, J.B., Katz, L.F., 2007. Experimental Analysis of Neighborhood Effects. Econometrica 75 (1), 83–119.

Lalive, R., Zweimüller, J., 2009. How Does Parental Leave Affect Fertility and Return to Work? Evidence from Two Natural Experiments. Quart. J. Econ. 124 (3), 1363–1402.

Lewis, J.D., Weigert, A., 1985. Trust as a Social Reality. Soc. Forces 63 (4), 967–985.

McEvily, B., Radzevick, J.R., Weber, R.A., 2012. Whom Do You Distrust and How Much Does it Cost? An Experiment on the Measurement of Trust. Games Econ. Behav. 74 (1), 285–298.

National Academies of Sciences, Engineering, and Medicine, 2015. The Integration of Immigrants into American Society. Washington, DC: The National Academies Press.

OECD, 2011. Naturalisation: A Passport for the Better Integration of Immigrants? Tech rep.. OECD Publishing, Paris.

Putnam, R.D., 2000. Bowling Alone: The Collapse and Revival of American Community. Simon and Schuster, New York.

Sapienza, P., Toldra-Simats, A., Zingales, L., 2013. Understanding Trust. Econ. J. 123 (573), 1313–1332.

Schönberg, U., Ludsteck, J., 2014. Expansions in Maternity Leave Coverage and Mothers Labor Market Outcomes after Childbirth. J. Lab. Econ. 32 (3), 469–505.

Selten, R., 1967. Die Strategiemethode zur Erforschung des eingeschränkt rationalen Verhaltens im Rahmen eines Oligopolexperiments. In: Sauermann, H. (Ed.), Beiträge zur experimentellen Wirtschaftsforschung. Mohr, Tübingen, pp. 136–168.

Steinhardt, M.F., 2012. Does Citizenship Matter? The Economic Impact of Naturalizations in Germany. Lab. Econ. 19 (6), 813–823.

Suárez-Orozco, C., Qin, D.B., 2006. Gendered Perspectives in Psychology: Immigrant Origin Youth. Int. Migr. Rev. 40 (1), 165–198.

Sutter, M., Kocher, M.G., 2007. Trust and Trustworthiness Across Different Age Groups. Games Econ. Behav. 59 (2), 364–382.

Sweetman, A., van Ours, J.C., 2014. Immigration: What About the Children and Grandchildren? In: Chiswick, B.R., Miller, P.W. (Eds.), Handbook of the Economics of International Migration, vol. 1B, 21. Elsevier, Amsterdam, pp. 1141–1193.

Uslaner, E.M., 2008. Corruption, Inequality, and the Rule of Law. Cambridge University Press, Cambridge.

Worbs, S., 2014. Bürger auf Zeit – Die Wahl der Staatsangehörigkeit im Kontext des deutschen Optionsmodells. Beiträge zu Migration und Integration, Band 7: Bundesministerium für Migration und Flüchtlinge.