Marginal Returns to Citizenship and Educational Performance

Christina Gathmann

LISER, University of Luxembourg and CEPR

Christina Vonnahme

RWI and Ruhr University Bochum

Anna Busse

Federal Ministry for Economic Affairs and Energy

Jongoh Kim

LISER

This Draft: October 8, 2021

Abstract

Citizenship is the most important right a host country can bestow on its immigrant population. Yet, little is known which citizenship policies work and who actually benefits from them. To answer these questions, we estimate the marginal returns to citizenship on children's school performance and skill development. For identification, we use two national reforms, which facilitated naturalization for first-generation immigrants and introduced birthright citizenship. We find substantial unobserved heterogeneity in returns with reverse selection on gains, i.e., the returns are highest for those with the lowest propensity of take-up. Citizenship significantly improves the school performance of immigrant children but has only modest effects on test scores. Policy simulations indicate that raising citizenship take-up would generate sizable benefits overall. Based on marginal treatment response functions, we also show that expanding birthright citizenship carries higher returns than facilitating naturalization.

Corresponding author: Christina Gathmann, Luxembourg Institute of Socio-Economic Research (LISER), University of Luxembourg and CEPR, christina.gathmann@liser.lu. Christina Vonnahme, RWI – Leibniz Institute for Economic Research and Ruhr University Bochum, christina.vonnahme@rwi-essen.de; Anna Busse, Federal Ministry for Economic Affairs and Energy, Berlin, Germany, anna.busse@web.de; Jongoh Kim, Luxembourg Institute of Socio-Economic Research (LISER), jongoh.kim@liser.lu.

We thank Martin Andresen, Simone Bertoli, Julia Bredtmann, Thomas Cornelissen, Fred Docquier, Hillel Rapoport, Kjell Salvanes, Joshua Shea, Steve Stillman, and participants at the Paris School of Economics, LISER Workshop on Migration, the CESifo conference on Education Economics, the UC Berkeley GSPP PhD Seminar, the 2nd CReAM/RWI Workshop on the Economics of Migration, the York Workshop on Labour and Family Economics (WOLFE), and the Verein für Socialpolitik for helpful comments and suggestions. The empirical analysis uses data from the National Educational Panel Study (NEPS), which was collected as part of the Framework Program for the Promotion of Empirical Educational Research funded by the German Federal Ministry of Education and Research (BMBF). Since 2014, NEPS is carried out by the Leibniz Institute for Educational Trajectories (LIfBi) at the University of Bamberg in cooperation with a nationwide network. We are grateful to the team at LIfBi for help with data access. Special thanks to Martin Andresen and Joshua Shea for answering our questions (Andresen, 2018; Torgovitsky and Shea, 2019).

1 Introduction

Citizenship is the most important privilege a host country can bestow on its immigrant population. Becoming a citizen and hence, a full member of the host society with all rights and responsibilities has obvious benefits in terms of political rights and participation; but do these benefits also reach beyond the political realm? Some think that the importance of citizenship for economic or social integration has declined as other titles like permanent residency, for instance, have expanded over time.[1]

Others, in contrast, argue that citizenship acts as an important catalyst for the economic and social position of immigrants and their descendants in the host country. They base their view on three arguments. First, citizenship improves the labor market position of immigrants through better pay and career options (Chiswick, 1978; Bratsberg et al., 2002; Gathmann and Keller, 2018).[2] A second argument why citizenship matters is that it reduces return intentions to the source country, which raises incentives to invest in skills that are highly valued in the host country – fostering educational, social and economic outcomes.(Dustmann, 2008).[3] A third argument is that citizenship shifts the attitude of natives and immigrants. Natives might have less biases or overt discrimination hampering the economic, social and educational integration of immigrants and their children.[4] Immigrants, in turn, may feel more welcome as equal members in the host society if offered citizenship. Citizenship could then change the role models immigrants and their children aspire to and raise parental efforts to invest in their children's human capital (Avitabile et al., 2014; Felfe et al., 2020).

Identifying the impact of citizenship faces sizable empirical challenges. Immigrant-native achievement gaps in school or the labor market are typically lower for traditional immigration countries like Canada or Australia than countries with restrictive citizenship policies like Germany or Sweden (Sweetman and Van Ours, 2014). While this comparison is suggestive, there are many other differences across countries that could explain the differential outcomes for immigrant children. Such differences are likely to influence immigrant selection, the environment immigrants find themselves in and hence, the choices immigrant families make. Access to citizenship sometimes varies within a country: children with one native and one foreign-born parent obtain citizenship automatically, for

[1] Shachar et al. (2017) provides a recent overview of the debate from different scholarly perspectives.

[2] Citizenship is often a prerequisite for certain government and other well-paid jobs; it also removes any restrictions on career and geographic mobility that immigrants face. These restrictions may apply to a wide range of occupations. In Germany, non-EU citizens have only had restricted access to regulated professions like lawyers, notaries, pharmacists or physicians prior to 2012 and could not easily change their occupation prior to 2005 or move to another EU member state, for instance.

[3] Moreover, employers might be more willing to invest in training an employee whose naturalization signals a long-term commitment to remain in the country (Lalonde and Topel, 1997).

[4] This argument relies on the assumption that discrimination is based on nationality and not on appearance or foreign-sounding names.

example. Yet, selection into mixed marriages is likely to create a different family environment than in families with two foreign-born parents.

To overcome these challenges, a recent literature has relied on large-scale reforms in citizenship law or peculiarities in the application process.[5] These studies focus on the reduced-form effect of eligibility on economic, educational and social integration outcomes. While valuable, intention-to-treat effects are not informative about who actually benefits from host country citizenship. Even more importantly, the reduced-form evidence remains silent on whether and how citizenship reforms could improve immigrant's outcomes. In particular, nothing is known whether immigrant children benefit more from birthright citizenship than from naturalization through their parents, for instance.

In this paper, we investigate the heterogeneity in the returns of citizenship for educational performance of immigrant children along both observable and unobservable dimensions. To do so, we rely on the marginal treatment effects (MTE) framework introduced by Björklund and Moffitt (1987) and generalized by Heckman and Vytlacil (1999, 2005, 2007) to model how the gains from citizenship depend on the unobserved costs of treatment. We then turn to policy simulations to learn more about the benefit of alternative citizenship policies. We first explore whether raising the overall take-up of citizenship could improve children´s school performance. We then use instrument-specific marginal treatment response functions (MTR) recently developed by Mogstad et al. (2018, 2020) to compare the gains of liberalizing naturalization with the gains of expanding birthright citizenship.

Our setting provides a unique opportunity to assess the heterogeneity in returns and the relative gains from alternative citizenship policies. First, we can exploit two national reforms, which substantially liberalized access to citizenship in Germany over the past decades. In both cases, eligibility for citizenship depends primarily on socio-economic characteristics like year of birth and years since migration. Most importantly, these demographic characteristics are not influenced by individual motivation or aspirations that may affect both the decision to naturalize and educational outcomes. Hence, we can use eligibility as instruments for the take-up decision and trace out its consequences for children's educational performance.

Second, the reforms cover both first- and second-generation immigrants. The reform in 1991 defined for the first time eligibility criteria for how first-generation immigrants could naturalize. Immigrants can either naturalize as teens when they reach age 16 and have lived in the country legally for eight years ('individual eligibility'); or they can naturalize as part of a family application for naturalization ('family eligibility'). In 2000, Germany further introduced birthright citizenship

[5] A first line of research estimates intention-to-treat effects of birthright citizenship on parental integration efforts (Avitabile et al., 2013), fertility behavior (Avitabile et al., 2014) and educational outcomes (Felfe et al., 2020). A second line of research investigates how eligibility for naturalization affects labor market performance (e.g. Bratsberg et al., 2002; Gathmann and Keller, 2018), marriage and fertility behavior (Gathmann et al., 2019) and social or political participation (Hainmueller et al., 2015, 2017) among first-generation immigrants.

for immigrant children. Children born in the country after January 1, 2000 with a foreign-born parent who has been living in the country legally for at least eight years obtain German citizenship at birth ('birthright eligibility').

How citizenship affects the integration outcomes of immigrants and especially their children is of enormous social, economic and political relevance. Migrant populations have grown rapidly in many countries, in particular among children and youth. By 2018, the shares of students with a migrant background have reached close to 50% in Australia or Switzerland, 40% in Canada and almost 30% in France, Germany, the UK and the United States (OECD, 2018b). At the same time, immigrant-native gaps in school subjects measured by language, math and science tests remain sizable (OECD, 2018a). Given the dynamic complementarities of learning for skill development later in life (Cunha and Heckman, 2007), these gaps have long-term effects on social and economic outcomes of children from immigrant families – thus cementing unequal opportunities and reducing inter-generational mobility. Identifying which citizenship policies foster integration becomes even more important in the face of current policy debates on citizenship. The US government has been discussing restrictions on birthright citizenship for children of undocumented or illegal parents, for instance. In Europe, the U.K. considered tightening the eligibility rules for naturalization, while Italy has debated to introduce birthright citizenship for children of immigrants.

Our results are as follows. We find strong evidence for reverse selection on unobservable gains. Returns to citizenship are higher for those least likely to take-up treatment. Citizenship leads to a decline in grade retention and an increase in high school attendance. We also see some improvements in school grades on key subjects like Math. Interestingly, we see few effects of citizenship on test scores administered by the survey team. That school performance improves but test scores do not could indicate that citizenship boosts non-cognitive skills like student motivation, which could manifest as less disruptive behavior or more participation in class. Another potential explanation is that teachers evaluate naturalized immigrant children more favorably than their non-naturalized peers.

The higher treatment effects for children with high unobserved costs compared to children with low costs is entirely due to improvements in school performance of children in the treated sample. Potential gains are essentially absent for children in the untreated state irrespective of their unobserved costs. Investigating the source of unobservable costs shows that families with high unobserved resistance to citizenship have a higher socio-economic status and are better integrated in the host country than families with low resistance. Citizenship is thus more of a complement to prior integration efforts and favorable family background when it comes to improving school performance among immigrant children.

We also find heterogeneity in the returns to citizenship along observable dimensions. Children originating from Central and Eastern Europe and the Former Soviet Union benefit the most from treatment: they improve their language and math skills and are more likely to attend high school. For immigrant children from Turkey, traditionally the largest immigrant group, the benefits are more muted. Yet, they are still more likely to attend high school if they hold host country citizenship. In contrast to immigrant origin, we see few gender differences in the benefits of citizenship.

The pattern of reverse selection on gains indicates that expanding citizenship eligibility or take-up could substantially improve the school performance of immigrant children. We start with estimating marginal policy-relevant treatment effects to assess the returns to raising the overall take-up of citizenship. We then turn to marginal treatment response functions in order to separate the potential gains of naturalization from the returns to birthright citizenship. In a final step, we simulate how reforming German citizenship law in the direction of a US-style model would affect take-up decisions in our data. Our policy simulations show that increasing take-up carries substantial gains in terms of better school performance. We further find that expanding birthright citizenship has larger benefits than facilitating naturalization.

We make several important contributions to the literature. Early studies in the citizenship literature used cross-sectional or panel data relating naturalization to labor market performance (Chiswick, 1978; Bratsberg et al., 2002). The more recent literature estimates reduced-form effects of eligibility (Felfe et al., 2020; Hainmueller et al., 2015, 2017) and how returns may vary across gender or country of origin, for instance (Gathmann and Keller, 2018; Gathmann et al., 2019). Our study makes three contributions to this literature: we estimate how actual citizenship acquisition affects school performance and skill development. Moreover, we explore who benefits from citizenship and how returns vary with the unobservable costs to take-up. Finally, we show which alternative citizenship policies for first- or second-generation immigrants benefit immigrant children and foster educational performance.

Our study also speaks to the literature on the assimilation and inter-generational mobility of immigrants in the educational system (Borjas, 1992, 1993; Dustmann and Glitz, 2011; Sweetman and Van Ours, 2014). Unlike existing work, we analyze how citizenship policies influence school performance and hence, inter-generational mobility of immigrants in the host society. Most importantly, we can explore who actually benefits from citizenship, and how the school performance of immigrant children might be boosted further by reforming existing citizenship policies.

We also contribute to an active literature that estimates marginal treatment effects. Most closely related are studies that focus on monetary returns to a college education (see e.g. Carneiro et al., 2011; Kaufmann, 2014; Nybom, 2017; Carneiro et al., 2017; Kamhöfer et al., 2018), secondary education

(Nybom, 2017) or early childcare (Cornelissen et al., 2018; Felfe and Lalive, 2018; Kline and Walters, 2016). Most papers document strong positive self-selection into the respective treatment based on net gains. We find the opposite pattern in our context: children with the highest gains are actually less likely to select into treatment. Several other public interventions have been analyzed including welfare (Moffitt, 2019) and disability insurance (Maestas et al., 2013; French and Song, 2014; Autor et al., 2019), discrimination among judges (Arnold et al., 2018), the impact of incarceration (Bhuller et al., 2020) or the performance of charter schools (Walters, 2018). Our study is the first to apply the framework to immigration and quantifying who benefits from citizenship.

Finally, we are one of the first studies using the marginal treatment response approach recently developed by Mogstad et al. (2018, 2020) to shed light on which citizenship policies work. Specifically, we use instrument-specific marginal treatment effects to inform policy-makers whether a reform of birthright citizenship carries higher gains than a reform of naturalization.

The paper proceeds as follows. We next discuss the two citizenship reforms in Section 2. Section 3 sets out the econometric framework to estimate marginal treatment effects and discusses the estimation strategy. Section 4 introduces the data source and the policy variation in our data. In Section 5 we present evidence that our identifying assumptions are satisfied and present our main results on selection into citizenship and the returns to citizenship on school achievement. Section 6 investigates which citizenship policy works and how alternative citizenship policies might affect educational performance. Finally, Section 7 concludes.

2 Background on Citizenship Reforms

Citizenship in Germany was traditionally tied to ancestry (*jus sanguinis*). Prior to 1990, there were no legal pathways to naturalize unless one could demonstrate German descent. Two national reforms in 1991 and 2000 overhauled the existing system and liberalized access to citizenship.[6]

The first reform, which came into effect on January 1, 1991, established rules for the naturalization of first-generation immigrants. The Alien Act (*Ausländergesetz (AuslG)*) required that adult immigrants (arriving after age 15) had to live legally in the country for fifteen years, while young teens (arriving between the ages of eight and fourteen) had to live in the country for eight years.[7] Adults who satisfied the criteria could include their spouse and dependent children under the age of

[6]The citizenship reforms were brought on the political agenda by three factors: first, the fall of the Iron Curtain and ensuing migration waves from Central/Eastern Europe and the Former Soviet Union. A second factor was the long-term presence of guest workers and their families without an opportunity to naturalize. And finally, a supreme court decision that turned down voting rights to foreigners in local elections, but called on the government to liberalize access to citizenship instead. Gathmann and Keller (2018) discuss the political process leading up to the reforms.

[7]If the applicant stayed abroad for no more than six months, the period of absence still counted toward the residency requirement. Longer stays abroad (between six months and one year) may still count for the residency requirement if they are shown to be temporary.

18 in the citizenship application. Children arriving under the age of eight or those born in the country could naturalize when they turned sixteen and satisfied a residency requirement of at least eight years. Hence, children and teenagers could naturalize either by satisfying the residency requirement themselves from age sixteen on ('individual eligibility') or if their parents included them in their application for naturalization ('family eligibility').

Beyond residency requirements, adult immigrants also have to demonstrate economic self-sufficiency, i.e., they should be able to support themselves and their dependents without welfare benefits or unemployment assistance. Teen immigrants arriving before the age of 15 had to complete a minimum of six years of schooling, of which at least four years have to be general education. Note that both requirements also need to be satisfied for a permanent work or residence permit. The applicants further had to renounce their previous citizenship unless they were EU citizens, must not be convicted of a severe criminal offense and declare their loyalty to the democratic principles of Germany's Basic Law.[8] Below, we define eligibility for citizenship using only objective rules on birth year or place and residency requirements. We do not incorporate economic self-sufficiency or the criminal record as these might themselves influence educational investments and school performance of immigrant children.

The Citizenship Act *(Staatsangehörigkeitsgesetz (StAG))* of 2000 introduced birthright citizenship. Children born in Germany in 2000 or later obtain German citizenship automatically if at least one foreign-born parent has lived in Germany legally for eight years (and had a permanent residency permit, available after five years of legal residence, for at least three years). Eligibility is assessed when parents register their newborn children, which is legally required.[9]

The 2000 reform also reduced the residency requirements for adult immigrants to eight years. The other requirements of the 1991 reform stayed the same: applicants could not have a criminal record, had to demonstrate economic self-sufficiency or basic education and their loyalty to democratic principles. In addition, the new law also required applicants to demonstrate adequate German language skills prior to naturalization. As before, the law of 2000 did not recognize dual citizenship in general though exemptions became more common in practice.[10]

[8]Children of bi-national marriages, for example, did not have to give up their dual citizenship until they turned 18. Exceptions were also granted if the country of origin prohibits the renunciation of citizenship or delayed it for reasons outside the power of the applicant, if the applicant was an acknowledged refugee or if the renunciation imposed special hardships on older applicants. Applicants with minor convictions, such as a suspended prison sentence up to 6 months (which would be abated at the end of the probation period), a fine not exceeding 180 days of income (calculated according to the net personal income of the individual), or corrective methods imposed by juvenile courts, were still eligible. Convictions exceeding these limits were considered on a case-by-case basis by the authorities.

[9]In most cases, children keep the citizenship of their parents. At age 21, children had to decide which citizenship to keep ('option model'). Since December of 2014, children who have grown up in Germany for at least 8 years and have finished at least 6 years of formal education can keep both passports.

[10]It became easier for older applicants and refugees to keep their previous citizenship. Applicants could also keep their nationality if it was legally impossible to renounce it or if it imposed a special hardship like excessive costs or serious economic disadvantages (e.g., problems with inheritances or owning property in their country of origin).

Table 1 provides an overview of the different pathways to citizenship and their main eligibility requirements defined by the 1991 and 2000 reforms: citizenship by birth for immigrant children born in Germany, citizenship through individual eligibility at age 16 and citizenship through family eligibility.[11] Figure 1 shows the pathways to citizenship in each category between 2000 and 2017. Across all age groups, the most important category is individual eligibility, which makes up for around 60% of all cases in Germany over this period. Birthright citizenship has also played an important role accounting for around 20% of cases. Naturalization as part of a family application makes up for around 10% of cases across all age groups. Among immigrant children, citizenship by birth and through family eligibility play a dominant role as many children in our sample have not yet reached the age threshold of 16 years. We next discuss our econometric framework to analyze who benefits from citizenship rights.

3 Econometric Framework

To investigate how host country citizenship affects school achievement and who potentially benefits in terms of observable, but also unobservable characteristics, we use the marginal treatment effects (MTE) framework developed by Björklund and Moffitt (1987), Heckman (1997) and Heckman and Vytlacil (1999, 2007).

3.1 Setup

Let Y_0 be a potential outcome like indicators of school achievement for children without host country citizenship and Y_1 the potential outcome for child i with host country citizenship; we refer to the latter as the treated state. Observed outcomes Y are related to potential outcomes through the switching regression model:

$$Y = NY_1 + (1 - N)Y_0$$

where $N = 1$ denotes the treatment. We model potential outcomes as linear in parameters:

$$Y_j = X\beta_j + U_j \qquad j = 0, 1 \tag{1}$$

where $E[U_j|X = x] = 0$, which is satisfied if we interpret equation (1) as a linear projection of Y_j onto X. The Xs denote child and parental characteristics that influence educational outcomes even independently of citizenship.

[11] The 2000 reform also stipulated a transitional rule. Between January 1 and December 31, 2000, immigrant parents who had lived in Germany for at least eight years could apply for naturalization of a child born between 1990 and 1999. As the rule was not widely known, take-up among potentially eligible children was below 10%; relative to the overall number of acquired citizenship, the transitional rule made up about 10% in 2000 and 2001 and less than 2% thereafter.

We follow the literature in modeling the take-up decision N, i.e., whether a child obtains host country citizenship, as a latent index model:

$$N^* = \mu_N(X, \tilde{Z}) - V \quad \text{with } N = 1 \quad \text{if} \quad N^* \geq = 0 \quad \text{and} \quad N = 0 \quad \text{otherwise} \tag{2}$$

where X denotes a matrix containing the observable characteristics and $\tilde{Z}$ the instruments, which are excluded from the outcome equation. In our case, the instruments are indicators equal to one if a child is eligible for citizenship through birthright, individual or family eligibility, respectively; and zero otherwise. The unobserved component of the selection equation in (2), V, enters negatively and hence, represents individual characteristics not observed by the econometrician that make a child less likely to obtain host country citizenship. Without loss of generality, we can transform the selection equation (2), $\mu_N(X, Z) - V \geq 0$, into $\mu_N(X, Z) \geq V$ and $\Phi(\mu_N(X, Z)) \geq \Phi(V)$ where Φ denotes the cdf of V. The term $\Phi(\mu_N(X, Z)) \equiv P(Z, X)$ is the propensity score, while the second term $\Phi(V) \equiv U_N$ denotes a random variable distributed in the unit interval.

Conditional on observables, U_N reflects any unobserved costs of applying for host country citizenship. Such costs could arise if parents deciding on the citizenship application for their child have a strong cultural attachment to their home country (Bisin and Verdier, 2011). Alternatively, immigrant families might face disadvantages in their home country if they give up their citizenship, like restrictions on buying property, for instance. Finally, the barriers might be a consequence of lack of information as parents might not be fully aware of the benefits of host country citizenship for their children.

We make the following two assumptions. Our first assumption is that (U_1, U_0, U_N) is independent of $(\tilde{Z}|X)$. In our context, this assumption implies that eligibility through one of the three pathways to citizenship is as good as randomly assigned and hence, independent of the unobserved gains and the unobserved resistance to treatment conditional on the control variables X. Second, we require a monotonicity assumption. The latent index model in equation (2) is equivalent to a monotonicity condition (Vytlacil, 2002) that can be weakened to a partial monotonicity condition in a setting with multiple instruments (Mogstad and Torgovitsky, 2018). The partial monotonicity assumption requires that for any two values of $\tilde{Z} = (z_l, z_{-l})$ and $\tilde{Z}' = (z_{l'}, z_{-l})$, $P(z_l, z_{-l}, x) \geq P(z_{l'}, z_{-l}, x)$ or $P(z_l, z_{-l}, x) \leq P(z_{l'}, z_{-l}, x)$ for all individuals. In our context, partial monotonicity implies that conditional on the other two eligibility channels and observables, each instrument weakly increases the likelihood of take-up. We discuss both assumptions and their empirical plausibility in more detail in Section 4.3.

The gain from treatment is given by $Y_1 - Y_0 = X(\beta_1 - \beta_0) + (U_1 - U_0)$. The first term represents how returns to citizenship vary with observable characteristics like gender or country of origin,

for example, in the treatment state, while the second term represents the unobserved gains from citizenship. Individuals can select into citizenship based on both observable and unobservable gains. In particular, the unobserved gains from citizenship $(U_1 - U_0)$ may be correlated with the unobserved costs of obtaining citizenship (U_N).

The marginal treatment effect (MTE) is defined by:

$$MTE(X = x, U_N = u_N) = E[Y_1 - Y_0 | X = x, U_N = u_N] \tag{3}$$

The MTE represents the gain from treatment for an individual with observed characteristics $X = x$ and unobservable characteristics $U_N = u_N$. Hence, the $MTE(X = x, U_N = p) = E[Y_1 - Y_0 | X = x, U_N = p]$ can be interpreted as the mean return to citizenship for individuals with $U_N = p$ who are just indifferent between naturalizing $(N = 1)$ and not naturalizing $(N = 0)$. Tracing the MTE at different U_N values reveals how the return to citizenship varies across the distribution of unobserved costs.

Under the conditional independence assumption we can trace the propensity score $P(Z, X)$ over the unit interval for a given X. In practice, we rarely have the data and variation in the instruments available, especially with a large set of control variables X and binary instruments as in our case. We therefore impose additive separability between X and (U_1, U_0) conditional on U_N (Carneiro et al., 2011; Brinch et al., 2017).[12] Based on this assumption, we can trace the MTE over the unconditional support of the propensity score $P(Z)$, as opposed to the support of $P(Z)$ conditional on $X = x$. A consequence of the additive separability assumption is that the shape of the marginal treatment effect does not depend on X except for its intercept. Under additive separability, we can rewrite the MTE as

$$MTE(X = x, U_N = u_N) = E[Y_1 - Y_0 | X = x, U_N = u_N] = x(\beta_1 - \beta_0) + K(p) \tag{4}$$

where $K(p) = E[U_1 - U_0 | U_N = u_N]$. The first term represents how the marginal gain of citizenship varies with observable characteristics, while the second term characterizes the heterogeneity in unobserved gains.

3.2 Empirical Specification and Estimation

The estimation proceeds in two steps. In the first stage, we estimate the selection equation, i.e., the empirical counterpart of equation (2). The dependent variable in the first stage is an indicator equal to one if a child has obtained German citizenship and zero otherwise. The instruments $\tilde{Z}$ are indicator variables for whether an immigrant child is eligible under the specific access option (by birthright,

[12] This assumption is slightly weaker than the full independence assumption, i.e., (U_1, U_0, U_N) is jointly independent of $(\tilde{Z}, X)$ that is typically imposed in empirical work (Carneiro and Lee, 2009; Carneiro et al., 2011; French and Song, 2014; Maestas et al., 2013).

through family or individual eligibility); and zero otherwise. For individual eligibility, we distinguish German- and foreign-born children because naturalization is an important pathway to citizenship for first-generation immigrants. It is much less important for second-generation immigrants who are often eligible for birthright citizenship.

As control variables X, we include child and family characteristics that influence take-up and educational outcomes. For the child, we include gender, a linear and squared term in age and the child's age of arrival.[13] We control for birth cohort fixed effects as the child's year of birth is an important determinant of eligibility. We further include a second order polynomial for parental years since migration to absorb general assimilation effects like improvements in parental labor market performance or language skills. We do not include parental education, employment or number of siblings. While these are potentially important determinants of family resources and parenting styles, they are themselves influenced by citizenship eligibility (Avitabile et al., 2014; Gathmann and Keller, 2018).

To adjust for heterogeneity in educational policies across states and time, we control for state fixed effects and year fixed effects. The year effects are allowed to differ across survey cohorts of children, which ensures that we only compare children´s outcomes within the same survey cohort and year. To adjust for differential propensities of take-up and school performance across source countries, we define five regions of origin: EU-15 member states, immigrants from Turkey, the Former Soviet Union (except the Baltic states), Central and Eastern European countries (including the Baltic states) and the rest of the world (Asia, Africa, the Middle East, North and South America, Australia and Oceania).

We estimate the first stage as a probit model thus assuming that the cdf of V, Φ, is a standard normal. From the estimates of the citizenship decision, we predict the propensity scores for treated and non-treated individuals. Recall that our second-stage estimation of the MTE is identified non-parametrically only over the support of $P(Z)$. Therefore, we impose common support of $P(Z)$ for treated and non-treated individuals. In addition, we also trim 1% of the observations with the thinnest common support in the data. We show in Section 5.5 below that our results are robust to alternative trimming margins.

In the second step, we estimate the outcome equation given by:

$$E[Y|X = x, U_N = u_N] = x\beta_0 + x(\beta_1 - \beta_0)p + K(p). \tag{5}$$

[13]Note that the child's age of arrival is equal to zero if the child was born in Germany.

The MTE is then calculated as the derivative of equation (5) with respect to p. We allow the observable gains to citizenship (the second term in equation (5)) to vary with child gender and, in some specifications, the region of origin.[14]

A crucial decision in economic selection models is the specification of $K(p)$, the unobserved gains to treatment (the third term in equation (5)). Our main results use a third-order polynomial in the propensity score, which restricts the MTE to be quadratic. In the robustness section, we demonstrate that more flexible, higher-order polynomials or semi-parametric methods to approximate $K(p)$ yield very similar results.

For the second stage, we use local instrumental variables (LIV), which estimates equation (5) and then calculates its derivative with respect to p. An alternative is the separate approach, which estimates the outcome equations in the treated and non-treated state separately and then calculates the MTE as the difference of treatment effects. As shown by Brinch et al. (2017), the separate approach may be particularly useful in the case of discrete instruments with limited variation. In our case, the instruments can take on six distinct combinations. Therefore, we can identify a flexible, high-order polynomial of the unobservable gains to treatment, $K(P(Z))$ even when using the LIV approach.[15] Our results are very similar if we use the separate estimation approach instead. We bootstrap all standard errors to account for the estimation error in the first stage.[16]

4 Data Source and Sample Description

4.1 National Educational Panel Study

To estimate the effect of citizenship on the educational performance of immigrant children, we rely on data from the National Educational Panel Study (NEPS). The panel collects detailed information on children's education and skill development (Blossfeld et al., 2011). The survey first samples schools and then randomly selects up to two classes in the target grade. All students in the selected classes are asked to participate and then followed over time even if they repeat or skip a grade.[17] To

[14] We include all other control variables from the first stage in the second stage but restrict their coefficients to be the same in the treated and untreated state.

[15] Note that the number of distinct instrument values $M = 6$ is lower than $2^3 = 8$ because birthright citizenship also implies parental eligibility. Suppose that the conditional expectations of the unobservables, U_1 and U_0, are specified as parametric functions linear in L parameters. Even without any control variables X, the LIV approach can identify a MTE with $L \leq (M-1) = 5$ parameters, while the separate estimation approach can identify $L \leq M = 6$ parameters (see Proposition 2 in Brinch et al., 2017). Hence, LIV can still identify $K(P(Z))$ as a polynomial of order six in our case.

[16] While the instruments vary at a more aggregated level than the individual child, each instrument varies at a different level (e.g., individual eligibility depends on birth cohort and years since migration, while birthright eligibility depends on parental years since migration and the child's birthplace). As a consequence, it is not obvious whether or how standard errors are to be clustered.

[17] The survey follows students who leave the sampled class or school. Nevertheless, tracking and testing school leavers and their families is more difficult and results in somewhat higher attrition rates than in the main sample.

select pre-school children, NEPS first samples primary schools and then surveys children in the day care centers that send children to the selected primary school. Our study covers the period from 2010 to 2018.

The NEPS data collects comprehensive information on children's school performance and their competencies as measured by standardized tests. Our main analysis focuses on grade retention, i.e., whether a student has to repeat a grade because of bad school grades in key subjects like Math or German, for instance. Grade retention is quite common in German schools. Approximately one out of six students repeats at least one grade during their school career; the share among foreign students is with one out of four even higher. Our second measure of school performance is whether the student attends the academic track in secondary school leading up to a high school degree after grade 12 or 13; or whether the student is sent to the non-academic track with graduation after grade 9 or 10. Traditionally, immigrant children have been much more likely to attend the non-academic track compared to their native peers.[18]

We also analyze school grades in Math and German. Grades are given by the teacher based on a child's performance in written and oral exams and class participation. Finally, we study a child's performance on standardized competence tests, which are rarely available in Germany for large samples of school-aged children.[19] It is important to stress that the tests were designed and administered by the research team and hence, need not be closely linked to the material covered in class. For the analysis, we analyze competencies in language and math, which are available for all children. The test for reading comprehension is similar to the PISA test and measures the ability to handle different types of texts encountered in daily life (like advertisements, literary or instructional texts). We use weighted maximum likelihood (WLE) scores with a mean of zero and standard deviation about one (in the first wave; values in subsequent waves slightly vary) to make effects comparable across subject and age. The online appendix provides a more detailed discussion of the tests and measures of school performance we use.

The NEPS data also contains information on parental education, working career and migration history, which is collected by interviewing the main caretaker, in 75% of the cases the mother.[20] We restrict the sample to children where the main caretaker is foreign-born though the parent might or

A-priori, it is unclear whether class or school leavers are positively or negatively selected compared to the sample stayers. We return to the question of selective attrition below.

[18]Parents, based on the recommendation by teachers in primary school, decide whether their child attends the academic track or not in most states. In a few states, the teacher recommendation has been mandatory (Baden-Württemberg until 2012, Bavaria, Brandenburg, Saxony, Saxony-Anhalt and Thuringia) but can be circumvented by parents requesting a slot or trial period in the academic track.

[19]Prior studies therefore rely on administrative data on school readiness, which is assessed by a pediatrician when the child is five years old (Cornelissen et al., 2018; Felfe and Lalive, 2018; Felfe et al., 2020).

[20]We have almost complete information on the main caretaker (with only 4% missing); information on the second parent or partner is scarcer with 23% missing. We return to this issue in the robustness section below.

might not have naturalized. The children might be born abroad or in Germany, i.e., are first- or second-generation immigrants.

Our analysis covers three cohorts of children spanning the period from kindergarten until the end of secondary schooling. The *child cohort* (born in 2005 and 2006) is on average nine years of age. They are still in kindergarten in the first wave, go through primary school and attend grades 5 and 6 at the end of our sample period. This cohort is either eligible for birthright citizenship or through their family. The *teen cohort* (born in 1999 and 2000) is on average 14 years of age. We follow this cohort from the entry into secondary school, which typically occurs in grade 5, to the end of compulsory schooling. This cohort is mostly eligible through birthright citizenship or their family. Some will also be eligible individually when they reach age 16. Finally, the *adolescent cohort* (born in 1995 and 1996) are on average 17 years of age. They attend grade 9 at the beginning of our sample period and have left secondary school at the end. This cohort is not eligible for birthright citizenship, but is either eligible through their family or individually. We follow the three cohorts until they reach the end of secondary school.[21]

Table 2 shows summary statistics of our sample of first- and second-generation of immigrants in the NEPS. Around 85% of children in our sample were born in Germany. Immigrant children born abroad came to Germany at the a mean age of four but as late as age sixteen.

4.2 Eligibility and Take-up of Citizenship

To study the effect of citizenship on children's school performance and skill development, we use the three access options to citizenship discussed in Section 2 as instruments for actual citizenship. Eligibility for any of the access options depends on observable characteristics: Eligibility for birthright citizenship varies with the child's birthplace and birth year as well as parental years since migration. Individual eligibility is primarily based on the child's birth year and the child's years since migration. Finally, family eligibility depends on parental years since migration, parental age and the child's birth year as only children under age 18 can be included in the citizenship application.

Eligibility is not perfectly correlated with actual citizenship for two reasons: the take-up decision of the immigrant and additional eligibility criteria.[22] All pathways to citizenship depend on residency requirements of the parent or child. While fulfilling the residency requirement is an important eligibility criteria, it is not the only one. Parents might fail to satisfy economic self-sufficiency or might have a criminal record, which would then disqualify them for naturalization, for example

[21] Our results are not sensitive to this restriction. We get very similar results if we only keep students until grade 10, which defines the end of compulsory schooling (see Section 5.5 below).

[22] There could also be measurement error in variables measuring citizenship or defining eligibility. In the case of classical measurement error, our estimates are biased toward zero making it less likely to detect a causal effect of citizenship on school achievement.

(see the discussion in Section 2). Furthermore, immigrants might not apply for German citizenship because they consider the costs higher than the perceived benefits of host country citizenship.

4.3 Monotonicity and Conditional Independence Assumptions

We now discuss the plausibility of our identifying assumptions and provide additional evidence supporting them. The model from Section 3 assumes partial monotonicity, which requires that the instruments weakly encourage take-up conditional on the observables X. Column (1) of Table A1 confirms that our eligibility variables weakly increase the likelihood of take-up. Individual eligibility is an important channel to obtain citizenship for foreign-born children, but not for German-born children. Most German-born children are eligible for citizenship by birthright or naturalize with their parents when they become eligible. Columns (2) to (4) further demonstrate that the correlation between citizenship and each instrument separately is positive with a similar heterogeneity for individual eligibility.

Our estimation approach further assumes that the instruments $\tilde{Z}$ are jointly independent of unobservable gains in the outcome equation (U_1, U_0) and resistance to treatment in the selection equation (U_N) conditional on X. Is this assumption plausible? It is important to note that the eligibility indicators, which serve as instrumental variables, depend only on observable demographic characteristics: the years since migration of the parent or child, the birthplace and birth year of the child (see the discussion in Section 4.2).

Yet, demographic characteristics like parental years since migration or children's birth year, might themselves be correlated with the unobservable gains (U_1, U_0). The parents of a child born in Germany have likely lived in the country longer and hence, might have more economic resources, better language skills or a different parenting style than parents of foreign-born children, for instance. We therefore control in all our specifications for the child's birthplace and parental years since migration (linearly and squared). The children's age and age of arrival, in turn, have a direct influence on skills like language development.[23] To control for these potential confounding influences, we include a full set of birth year fixed effects and the child's age (linearly and squared) as controls in our baseline in equation (5).

As eligibility only depends on observable demographic characteristics, the instruments are by definition uncorrelated with any individual motivations and aspirations to succeed in the source country. These aspirations will be captured by the unobserved barriers or costs to treatment (U_N). Our instruments are also independent of requirements for economic self-sufficiency or minimum

[23]It is well known that learning another language becomes more difficult after around age eleven. Further, there could be cohort effects in school achievement if schools or teachers become better overall or in teaching classes from diverse backgrounds, for instance.

years of schooling, which might influence the unobserved educational outcomes (U_1, U_0); the reason is that our eligibility indicators do *not* depend on these additional eligibility criteria while the actual treatment citizenship does.

The conditional independence assumption could still be violated for at least three reasons. First, the citizenship reforms might have changed the selection of immigrants over time. In Germany, more immigrants from the Former Soviet Union or Central and Eastern Europe entered the country after 1990 (Gathmann and Monscheuer, 2021), while the number of EU immigrants has increased over the past decade. Recent immigrants also tend to be more educated, which is likely to influence the educational performance of their children. To address this concern, our baseline specification controls for the geographical composition of immigrants by including region of origin fixed effects distinguishing between immigrants from five different groups: immigrants from traditional EU member states (like Greece, Italy or Spain), immigrants from Central and Eastern Europe (like the Czech Republic or Poland), Turkish immigrants, immigrants from countries that used to be part of the former Soviet Union (like Kazakhstan or Latvia) and all others (lumping together immigrants from Africa, Asia and America).

A second concern is selective fertility after the 1991 and 2000 reforms. Gathmann et al. (2019) find that citizenship for adult immigrants postpones marriage and childbearing, while Avitabile et al. (2014) show that the birthright citizenship reduced fertility but fostered additional parental investments into the children born. As parental investments are often complements to skills learned in schools, these children might have higher unobservable gains than children not eligible for birthright citizenship. To take these differences into account, we include fixed effects for the child's birth year and birthplace. We also consider birth month dummies and dropping children born after the 2000 reform to control for any fertility adjustments in Section 5.5.

A third concern could be selective return migration. Our estimation compares the returns to citizenship for eligible and non-eligible children of immigrants conditional on them staying in Germany. We argue that identifying the returns to citizenship among stayers, rather than among all immigrants entering the country at some point, is an important policy-relevant treatment effect of interest. Selective return migration or sample attrition more broadly could still influence our results if it is correlated with eligibility. Access to citizenship, for instance, might reduce return migration to the home country. Suppose there is positive selection into return migration so the children with the best skills or school performance are more likely to return to their home country or drop out of the sample for other reasons. In that case, citizenship eligibility, by reducing return migration, will keep more children with better skills in the sample; we will then obtain an upward bias in our estimated returns to citizenship. To investigate the relationship between attrition and our instruments, we

analyze two types of attrition: temporary dropout and permanent attrition. Temporary dropout, i.e., a student did not participate in a wave because of sickness or absence from school is very common in our sample (36%). Permanent drop-out, which occurs if the student leaves the survey permanently and could not be tracked further, is with 6% very rare. Appendix table A2 shows that none of our eligibility indicators is correlated with permanent dropout from the sample (see Column (2)). As such, there is no systematic bias in our sample because of long-term attrition. Temporary dropout for one wave, in turn, is higher for students eligible for birthright citizenship (see Column (1)). The effect is with 4.6 percentage points (a 12% increase) rather modest, however, compared to the much higher temporary dropout among older students. Most importantly, temporary dropout should not systematically bias our results as we still have information on school performance for these students in our panel.

Finally, one might worry that eligibility can create spillover effects on the child's performance or between children even independently of actual citizenship. In the first case, eligible students improve their performance because eligibility makes them feel welcome in the host society and boosts their motivation to succeed in school, for example. We would then observe better school performance or skill development even if the eligible student does not naturalize. In the second case, eligible students might have a positive influence on the school performance or skill development of their non-eligible peers. Under both scenarios, eligibility would boost the outcomes of non-treated individuals and hence, produce MTE estimates that are lower bounds of the true effects.

5 Empirical Results

5.1 Selection Into Citizenship

We start out with the estimates of the selection equation (2). Table 3 shows marginal effects from a probit model where the dependent variable is equal to one if the child holds German citizenship and zero otherwise. The first two columns show the results for the overall sample, while the last two columns report estimates for the sub-sample for which we have information on our main outcome grade retention. The first specification distinguishes between immigrants from Europe and outside Europe, while the second specification shows differential effects for the five main origin regions.

The results in Columns (1) and (2) demonstrate that the three access options have a strong effect on the likelihood of obtaining citizenship. Birthright citizenship, which is available for children born in Germany, raises the probability of having a German passport by 13.4 percentage points in the full sample. Compared to children born in Germany, foreign-born children are 11.7 percentage points

more likely to naturalize under individual eligibility.[24] Family eligibility also raises the likelihood of host country citizenship by up to 15 percentage points. The Chi^2-statistics at the bottom of Table 3 confirms that the three access options have a strong effect on selecting into citizenship across all specifications.

We find few gender difference in the take-up of citizenship; if anything, girls are slightly more likely to obtain host country citizenship. Foreign-born children, in contrast, are much less likely to obtain citizenship than German-born children with foreign-born parents. This lower take-up is not surprising because foreign-born children have to apply for naturalization while many German-born children obtain citizenship by birthright. Take-up rates vary substantially across immigrant origins: immigrants from Europe as a whole are about 4 percentage points more likely to naturalize than non-European immigrants (see Column (1)). Column (2) shows that nationals from traditional EU member states (like Italy or Greece, for instance) are less likely to naturalize than immigrants from the rest of the world. The lower take-up among immigrants from traditional EU member states should not be surprising as the Single Market grants EU nationals very similar rights than natives reducing the gains compared to immigrants from countries not part of the EU. Column (2) further shows that immigrants from countries that were part of the Former Soviet Union are much more likely, while children of Turkish origin have a 16 percentage point lower likelihood to take-up German citizenship than immigrants from the rest of the world. We also see a strong influence of age: older immigrants and immigrants who arrived in the country at an older age have lower take-up rates than younger children and children who arrived at a young age or were born in Germany.

Figure 2 plots the predicted propensity score $\widehat{P(Z,X)}$ separately for treated and untreated immigrant children. The figure reveals that we have common support for the full sample over the range from 0.42 to 0.99.[25] The less than full support implies that we cannot evaluate the effects of citizenship for immigrants with very low unobserved costs – unless we were willing to extrapolate based on functional form assumptions like normality, for instance. The common support covers immigrants with medium to high unobserved resistance to take-up. We believe this range covers an empirically highly relevant margin as take-up rates of citizenship range from 35% to 80% across destination countries. Furthermore, immigrants with medium to high unobserved costs is the relevant group to answer important policy-relevant questions, e.g. who might benefit from expanding access to citizenship or how take-up can be increased.[26] Below, we aggregate the MTEs into common

[24] Individual eligibility, which is available for teens starting at age 16 plays a more important role for the sub-sample with information on grade retention because grade retention is basically unknown in pre-school and primary school.

[25] Appendix figure A1 shows that for the grade retention sub-sample, the common support in the propensity score ranges from 0.36 to 0.98.

[26] We are not the only study estimating heterogeneous treatment effects with limited common support in the propensity score: Maestas et al. (2013) obtain coverage between 0.57 and 0.85 and French and Song (2014) cover the range between 0.45 and 0.85 for disability receipt, for instance.

treatment effects like ATE, ATT, ATU and LATE by re-scaling the weights to integrate to one over the region of common support following Carneiro et al. (2011).

5.2 Marginal Treatment Effects with Unobservable and Observable Heterogeneity

We now turn to the question whether and how citizenship affects educational outcomes. We focus here on grade retention as it is an important indicator of school performance, in particular for weaker students. We discuss the impact of citizenship on other outcomes like school grades and test scores in Section 5.4.

Figure 3 plots the marginal treatment effect evaluated at the mean of observable characteristics X, which relates the unobservable gains from treatment $U_1 - U_0$ (on the y-axis) to the unobserved costs of obtaining citizenship U_N (on the x-axis) for immigrants just indifferent between being treated and not treated. Panel (a) plots the MTE obtained from the local instrumental variable approach, Panel (b) the MTE obtained from the separate approach. Both panels show a downward sloping marginal treatment effect curve: individuals who face higher unobserved costs to citizenship would be less likely to repeat a grade in school if they obtained citizenship than children with low unobservable costs. The MTE curve starts out flat at zero implying that there is no effect on grade retention for children with low resistance to treatment. Hence, immigrant children who are most likely to obtain citizenship in Germany do not improve their school performance as measured by grade retention. Children with high resistance, in turn, would substantially reduce their likelihood of grade retention if they obtained citizenship. In particular, immigrant children with an unobserved resistance larger than $U_N > 0.8$ have a statistically significant lower probability to repeat a grade if they choose treatment.

Hence, the MTE curve indicates a reverse selection on unobservable gains – returns to citizenship in terms of grade retention are higher for those least likely to take-up citizenship.[27] There are several potential explanations for this finding: immigrants families either do not know about the potential benefits of citizenship for their children. Alternatively, the decision to take-up citizenship is influenced by family considerations that weigh more heavily than improvements in children's school performance. We explore potential explanations for this pattern in Section 5.6 and study its implications for policy in Section 6 below.

Table 4 shows the coefficients of the second-stage estimation based on equation (5). The estimates confirm the presence of essential heterogeneity: the coefficients on the linear and squared propensity

[27] We are not the only study finding reverse selection on gains for Germany. Cornelissen et al. (2018) observe that minority children benefit more from early childcare but are less likely to attend than native children.

score are statistically significant and the p-values shown at the bottom of the table indicate that returns to citizenship vary systematically with the unobserved costs of take-up. Once we allow for different effects by region of origin (in Column (2)), we also see some heterogeneity in returns to citizenship along observable characteristics as shown by a p-value of 0.049 for observable heterogeneity.

In the untreated state (corresponding to β_0 in equation (5)), immigrant children from non-European countries and from Central and Eastern Europe are less likely to repeat a grade, while immigrants from the Former Soviet Union and Turkey are much more likely to repeat a grade. The effect turns around with treatment, shown as the coefficient on the interaction between the propensity score and region of origin and corresponds to $(\beta_1 - \beta_0)$ in equation (5). Immigrants from the Former Soviet Union and Turkey are less likely to repeat at grade if they naturalized than immigrants from the rest of the world.

Do we see any selection of immigrants from different origins into treatment based on the unobservable gains? That is, do we observe that $\gamma * (\beta_1 - \beta_0) > 0$, which would indicate a positive selection on gains along observable characteristics? A comparison of Tables 3 and 4 reveal that immigrants from the Former Soviet Union gain from treatment and have higher take-up rates than immigrants from the rest of the world. There is also a consistent pattern for immigrants from Central and Eastern Europe: they have a lower take-up rate and do not seem to gain from treatment in terms of grade retention relative to immigrants from the rest of the world. The pattern for Turkish immigrants is not consistent with positive selection on gains: Turkish children do gain from treatment but are much less inclined to take-up citizenship. Hence, immigrant families from Turkey seem to face higher unobserved barriers to citizenship than other immigrant families.

5.3 Aggregate Treatment Effects

So far, we have discussed how the returns to citizenship vary along observable and unobservable characteristics. To learn more about the overall returns to citizenship, we derive summary measures of treatment effects. The average treatment effect (ATE), the treatment effect on the treated (ATT), the treatment effect on the untreated (ATU) and the local average treatment effect (LATE) are obtained by aggregating the MTE curve over the relevant range of observable and unobservable characteristics. The ATE aggregates over the observable characteristics (X) and unobservable costs (U_N) of all immigrant children. The ATT aggregates over the characteristics of children with German citizenship, while the ATU aggregates over the observable and unobservable characteristics

of immigrant children without citizenship. The LATE is obtained by aggregating over immigrant families who switch treatment status, i.e., the complier population.[28]

Column (1) of Table 5 shows that the ATE for grade retention is –0.142. Hence, if we pick a random immigrant child from the population, the ATE suggests that citizenship reduces their risk of grade retention in secondary school by 14.2 percentage points or 40% of a standard deviation. Citizenship thus enables immigrant children to boost their school performance and substantially reduce their risk of repeating a grade. In line with our finding of reverse selection on gains in Figure 3, the ATT for grade retention is –0.077 and not statistically significant. The ATU, in turn, is –0.359 and statistically significant. The returns to citizenship in terms of grade retention are much stronger in the sample of untreated children where one out of four children repeats a grade. The ATU indicates that these students who have the highest risk of grade retention would benefit the most from citizenship.

To illustrate how different treatment effects aggregate over observables and unobservables, Panel (a) of Figure 4 shows the weights of the ATT and the ATU against the unobserved costs U_N. As the ATT (shown as red crosses) aggregates over the treated sample, it puts much more weight on individuals with low unobservable costs to citizenship ($U_N < 0.7$). Weights of the ATT are low and approach zero for individuals with high resistance to citizenship ($U_N > 0.7$). The opposite is true for the ATU (shown as blue crosses): here, weights are highest for individuals with very high costs to obtaining citizenship (U_N close to 1), decrease for individuals high resistance ($U_N > 0.6$) and approach zero for individuals with medium to low costs to citizenship ($U_N < 0.6$).

Panel (a) of Figure 4 also plots the MTE if evaluated at the average characteristics in the sample (MTE), which is the same as in Figure 3 above. In addition, we plot the MTE if evaluated at the observable characteristics of the treated (MTE(ATT)) or the characteristics of the untreated (MTE(ATU)). By construction, the three MTE curves have the same shape because the additive separability assumption implies that observable characteristics only influence the location of the MTE curve but not its shape. Yet, the three MTE have also the same location irrespective of whether we evaluate them at the mean in the sample (the black curve), the characteristics of the treated (the red curve) or the characteristics of the untreated (the blue curve). That pattern indicates the treated sample has overall quite similar observable characteristics than the untreated sample.

We can further explore how our treatment effects compare to the local average treatment effect (LATE). The LATE can be obtained from the MTE by aggregating the observable and unobservable characteristics over the complier population, i.e., those immigrant children who change treatment

[28]Table 2b in Heckman and Vytlacil (2007) and Table 1 in Mogstad et al. (2018) provide a formal representation of the weights for the different treatment effects.

status in response to one of the instruments. Table 5 shows that the LATE is –0.092 and not statistically significant just like the ATT. Panel (b) of Figure 4 plots the weights for the LATE estimator (the red crosses) as well as the MTE curve evaluated at the observable characteristics of the complier population (the red curve). LATE gives higher weights to individuals who are more affected by the instrument conditional on the control variables. Figure 4 shows that these are individuals with resistance in the range of $0.65 \leq U_N \leq 0.85$. In contrast, LATE, like the ATT, puts low weights to children with the very low and very high resistance to citizenship. Moreover, the MTE curve (shown in red) evaluated at the X of the complier population lies above the MTE curve for the average X in the sample (the black curve) indicating that the complier population has observable characteristics that are associated with only small benefits to citizenship. Based on the LATE estimates, we would wrongly come to the conclusion that there are few returns to citizenship for school children. Such inference would not only be misleading but also completely miss the sizable potential gains from citizenship among the untreated.

5.4 Evidence on Track Choice, School Grades and Test Scores

So far, we have documented that immigrant children may benefit from citizenship by reducing their probability of repeating a grade in school. The decision to pass or repeat a grade is a combination of the performance of the student and an assessment of the student's skill and behavior by the teachers. Moreover, it is an outcome that matters mostly for weaker students. Do we see a similar pattern of reverse selection on gains for other educational outcomes like high school attendance, school grades and test scores?

To explore this question, we re-estimate our baseline model in equations (2) and (5) where the dependent variables are whether a child attends the academic track (leading to a high school degree), school grades (in Math and German) and cognitive tests (in Math and Reading comprehension). Table 5 shows the aggregate treatment effects we obtain by aggregating the MTE curve over the relevant range of X and U_N.

There are several noteworthy patterns. We observe strong evidence for a reverse selection on gains for track choice, just as we did for grade retention (see Column (2)). Hence, the average treatment effect on the untreated (ATU) indicates that treatment could raise the likelihood of immigrants attending the academic track by 15.4 percentage points. Given a standard deviation of 0.48 in the sample, the ATU implies an increase of about 30% of a standard deviation if untreated

children would obtain citizenship. The ATE and ATT, in turn, indicate no effect of citizenship on the type of secondary school whatsoever.[29]

We also find some evidence for a reverse selection on gains for school grades. Lower grades indicate better performance with one being the best and six the worst grade. The ATU for both German and Math grades (see Columns (3) and (5)) show that untreated immigrant children would benefit from citizenship through better grades in those key subjects. The ATU for Math grades suggests a statistically significant improvement of about –0.324 or a third of a grade. Given average school grades in Math of 2.77, obtaining citizenship could improve Math grades of the untreated by 12%. These gains would roughly close the gap in Math grades between immigrant and native children. The pattern is weaker and not statistically significant for grades in German though.

We see few effects of citizenship on actual test scores in Math and German (see Columns (4) and (6)). For test scores in Math, the ATU is larger than the ATT suggesting reverse selection on gains though the effects are not statistically significant. We do not see a reverse selection on gains for reading comprehension.[30] It is important to stress however, that the standardized tests administered to the children are not part of the school curricula but designed by the survey team to test children's learning achievement. As such, the material in the tests need not be closely related to the material covered in class when the test was taken. The p-values of 0.817 and 0.121 at the bottom of Table 5 further indicate that the unobserved costs of take-up are not strongly related to test performance conditional on observable characteristics. Hence, immigrant children with higher costs to citizenship do not systematically perform worse or better on standardized tests than immigrant children with low costs to citizenship.[31]

Across all measures of educational performance, we see that citizenship improves outcomes especially for weaker students who are most likely to repeat a grade or to attend the non-academic track. The evidence on reverse selection on gains indicates that there could be sizable gains for increasing take-up of citizenship among immigrants with high unobserved costs. We find weaker effects for grades and test scores, which reflect performance across the whole grade and ability spectrum. It appears that citizenship has little direct influence on core skills; instead, citizenship

[29]Track choices are typically made by the parents based on the recommendations of primary school teachers. Supplementary evidence indicates that teachers seem to be more likely to recommend immigrant children to the academic track if they were naturalized.

[30]One potential explanation for the null effect in reading comprehension is that immigrants who apply for naturalization have to demonstrate basic German language skills. The language test is required of young immigrants if they apply for citizenship at age 16 or later (through individual eligibility). If the family applies for citizenship, the requirement only applies to the parent applying; yet, the better German skills of the parent could also benefit the child's language skills.

[31]The absence of essential heterogeneity implies that, under additive separability, equation (4) becomes $E(Y_1 - Y_0|X = x, U_N = u_N) = E(Y_1 - Y_0|X = x)$, so $Y_1 - Y_0$ is mean independent of U_N given $X = x$ and all the treatment parameters are the same over our common support (so, MTE = ATE = ATT = ATU = LATE). Note that there still could be essential heterogeneity in gains outside the common support for individuals with very low U_N.

seems to influence how teachers assess the performance of immigrant students. One reason could be improvements in non-cognitive skills like student motivation: immigrant students who are citizens of the host society might be less disruptive and be more active in class. A change in student behavior in the classroom could affect school grades but need not necessarily manifest in better test scores. Another explanation is that there is an implicit or explicit bias in how teachers view immigrant students, which is reduced with citizenship. Anecdotal evidence indicates that immigrant children are less likely to be recommended for the academic track, for instance. Children with German citizenship might therefore be viewed more favorably by teachers, which would explain why their school performance would improve despite few visible test score gains.[32]

The p-values at the bottom of Table 5 further reveal that there is sizable observable heterogeneity for all outcomes with the exception of German language skills. Appendix table A3 shows the estimates of equation (5) for our additional outcome variables. We again see few differences in returns to citizenship across gender. Students from most origin regions benefit from citizenship in terms of a higher likelihood of attending the academic track leading up to a high school degree. We also find that students originating from Europe (traditional EU and Central and Eastern Europe) improve both their grades and test scores with citizenship. For immigrant children from Turkey, we find only modest gains from citizenship in terms of grades and test scores.

5.5 Specification Checks

We now demonstrate that the reverse selection on unobservable gains to citizenship is very robust to alternative assumptions and samples. We focus hereby on our main outcome grade retention.

We first probe the sensitivity of our results to alternative choices of the empirical model. The specification of the polynomial in the propensity score $K(P(Z))$ is key to model the correlation between unobserved gains and unobserved barriers. Our baseline results in Table 3 use a third-order polynomial to approximate $K(P(Z))$, which implies a quadratic MTE function. A quadratic MTE function might not be flexible enough to capture unobserved heterogeneity, especially since we find the largest gains for those with high unobserved costs. Panel (a) of Figure 5 shows the MTE curve if we use a third-order and fourth-order MTE curve instead. The MTE curve looks almost identical to the MTE curve in the baseline. As such, a quadratic MTE curve captures the correlation between unobserved gains and unobserved costs reasonably well. Rather than using polynomials, we estimate a more flexible semi-parametric approach using double residual regression instead.[33] The result

[32]We cannot distinguish with our data whether teachers assess non-citizens worse than native students or whether they actually evaluate immigrant students with citizenship more favorably than native students.

[33]In the first step, we obtain residuals from regressions of Y, X and Xp on polynomial in p; in the second step, we estimate β_0 and $\beta_1 - \beta_0$ by regressing the residualized outcome on the residualized regressors. We then use a local

shown in Column (1) of Appendix table A4 finds a very similar pattern: the treatment effect for the untreated exceeds the treatment effect of the treated implying reverse selection on gains. Instead of the local instrumental variable approach, which calculates the derivative of the outcome equation to obtain the MTE, we estimate the outcome equation (5) separately for the treated and untreated state and then take the difference. Column (2) of Appendix table A4 shows that the separate approach not only yields a similar MTE curve (see Panel (b) of Figure 3), but also generates very similar treatment effects as the local IV approach: an ATU that is larger in absolute value than the ATT while ATE and LATE estimates lie in between. Finally, our baseline trims 1% of the observations with the thinnest common support. Panel (b) of Figure 5 shows that we obtain a similar MTE curve with a trimming margin of 2% and a slightly flatter one if we trim 5% of the observations.

Our main results further show that immigrant children with citizenship are more likely to attend high school. One might therefore worry that tougher requirements in the academic track might result in more grade retention or worse grades and hence, bias our treatment effects downward. Column (3) of Appendix table A4 re-estimates our baseline adding a control for track choice. The resulting ATE and ATT are very similar to the baseline, while the ATU becomes even larger indicating that citizenship carries even higher benefits for the untreated if we also account for the school track attended. It is well known that small age differences might have large effects on children's school performance. As age is also correlated with eligibility for citizenship, our estimates might reflect age differences. To check for the influence of small age differences, Column (4) in Appendix table A4 adds birth month dummies in addition to year of birth dummies, child age and age squared to the specification, which leaves the results unchanged.

Parents might have adjusted their fertility behavior in response to the citizenship reforms. In particular, parents may have delayed conception or birth to ensure that their child is eligible for birthright citizenship. If parents delay conception, we should see a spike in birth in January of 2000 and a (corresponding) decline of births in December of 1999. Vital statistics however, exhibit no discontinuity in the number of immigrant children born in the time period before or after January 1, 2000 (see Felfe et al., 2020). Parents scheduled to deliver around the cutoff date could have tried to delay the birth of their child until January 1, 2000. Yet, Column (5) of Appendix table A4 shows that dropping children born in December of 1999 and January of 2000 does not affect our treatment effects. Parents might also reduce their preferred number of children, have children at later age and invest more in them (Avitabile et al., 2014; Gathmann et al., 2019) – in line with the quantity-quality trade-off (Becker and Lewis, 1973; Becker and Tomes, 1976). Column (6) drops all children born in

polynomial regression of the residualized outcomes on p to obtain an estimate of $K(p)$. Finally, the MTE is calculated by taking the derivative of $K(p)$.

April 2000 or later. While this restriction reduces our sample by almost 50%, the treatment effects exhibit a very similar size and pattern.

Finally, we check whether our results are robust to alternative definitions of our sample. As the selection of immigrants arriving in Germany has changed over time, the type of immigrants obtaining citizenship might have changed as well. In Column (7) of Appendix table A4, we drop all guest workers, i.e., immigrants arriving in Germany prior to 1975. Interestingly, this has very little impact on the estimated treatment effects suggesting that benefits of citizenship exist for early and later arrivals alike. Our main results include all secondary grades; yet, some immigrant children leave school after the end of compulsory schooling (grade 9 or 10). To see whether that influences our results, we drop all grades after the end of compulsory schooling (in Column (8) of Appendix table A4). This restriction has no effect on the results suggesting that benefits of citizenship apply to all grades in school. In Column (9) of Appendix table A4, we further drop immigrant children aged 18 or older as they could obtain citizenship through marriage. These are few cases and dropping them has no effect on our results. To address the concern about missing parental information, we also drop all immigrant children where we do not have information on the second parent (in Column (10) of Appendix table A4). These include all single-parent families but could also include some families where the second parent could be a native. The treatment effects we obtain from the three alternative samples are very similar to the baseline in Table 5. Finally, we drop potential ethnic Germans from our sample in Column (11), who had faster access to German citizenship than other immigrant groups. The results are again very similar to our baseline.

Taken together, the alternative estimation approaches, specifications and sample restrictions do not change our main result that citizenship has sizable benefits for those immigrant children who are least likely to take up citizenship.

5.6 Which Immigrants Have High Unobserved Costs to Citizenship?

The fact that immigrant children with the highest unobserved benefits also face high costs to obtain citizenship begs the question what factors might explain the reverse selection on gains. To see whether the unobserved gains $E[U_1 - U_0|U_N = u_N]$ stem primarily from treated or untreated immigrant children, Figure 6 plots the unobserved gains against the unobserved costs of citizenship separately for treated, $E[U_1|U_N = u_N]$, and untreated immigrant children, $E[U_0|U_N = u_N]$. Essentially all of the unobserved gains we observe are accounted for by children who obtained citizenship. Among untreated children, potential gains from citizenship in terms of school performance are almost flat across the distribution of unobserved costs. In the treated sample, in contrast, unobserved gains in

school performance are sizable for immigrant children with high barriers to treatment, but small or even absent for children with very low costs to citizenship.

To gain some insights into the nature of these unobserved costs U_N, we turn to the family characteristics available in our data. We plot family background characteristics, which are obtained as residuals after taking out of our full set of control variables and instruments, against the unobserved costs U_N separately for treated and untreated immigrant children.

Unobserved costs could result from immigrant families lacking information or underestimating the benefits of citizenship for their children. In the absence of direct information, we use maternal education and the number of books at home as a proxy. Parents with more education are more likely to have better information or know better where to find it. Panels (a) and (b) of Figure 7 thus plot the share of mothers with more than primary education and the number of books in the family home. Immigrant families with high costs are more educated and possess more books than families with low unobserved costs. Lack of information or knowledge does therefore not seem to be the main explanation for high unobserved costs.

Families might resist host country citizenship because the have more traditional family values in which mothers do not work and take care of children in the home. These values might clash with the diversity of family life in the host country. To assess this explanation, Panels (c) and (d) of Figure 7 plot the share of working mothers and the number of siblings in the family. Immigrant parents with low unobserved costs have larger families and lower maternal employment than families with high unobserved costs. If anything then, families with low unobserved costs seem to have a lower socio-economic status and more traditional family values than families with high unobserved costs. Hence, the unobserved costs do not seem attributable to a clash of values between the home and host country.

Finally, parents might have a preference that their child is socialized in the culture of their home country rather than becoming fully integrated in Germany (Bisin and Verdier, 2011). As a proxy, we study whether the family speaks German at home or whether the child feels foreign in Germany. Panels (e) and (f) of Figure 7 show that families with high unobserved costs are more likely to speak German and feel less foreign in Germany than families with low unobserved costs.

The evidence on family background, unobserved costs and gains allow three conclusions: first, the higher treatment effects for children from families with unobserved costs is almost entirely due to improvements in school performance among treated children. Immigrant families with high unobserved costs and high potential gains for their children are more educated, have higher maternal employment rates and seem better integrated into the host society than immigrant families with low costs. Second, citizenship improves the school performance of children coming from families

with more favorable socio-economic background and prior integration efforts. As such, citizenship should be viewed as a complement to the origin family and previous integration efforts; it does little to counteract or compensate for less favorable socio-economic family backgrounds or lack of integration. Finally, we see small differences in family background between treated and untreated children conditional on unobserved costs.

6 Policy Simulations

The fact that immigrant children with high unobserved costs have the largest gains from citizenship raises the question whether a reform of citizenship policies could further improve the educational trajectories of immigrant children.

6.1 Raise Take-up of Citizenship Overall

We first use our econometric framework in Section 3 to simulate alternative citizenship policies. We consider counterfactual policies that increase $P(Z)$, the propensity to take up citizenship, but do not affect potential test scores or the unobservables determining selection (U_1, U_0 and U_N). Hence, knowing how the policy changes the distribution of the perturbed propensity score is sufficient to calculate a policy-relevant treatment effect (PRTE). Because we do not identify the propensity score over the full unit interval, we estimate marginal policy-relevant treatment effects considering small changes following the approaches in Heckman and Vytlacil (2001, 2005) and Carneiro et al. (2011).

We implement the alternative policies in two ways: adding a small amount α to the propensity score or increasing the propensity score by $\alpha\%$ where α can be arbitrarily small. Denote the alternative distribution of the propensity score by P^α. We then use the MTE from Section 3 weighted over the set of people shifting into treatment under the alternative policy P^α to obtain the marginal policy-relevant treatment effect.[34]

Row (A) of Table 6 shows that encouraging take-up among immigrant children carries sizable gains. Grade retention among immigrant children would decrease by around 6 percentage points even for small changes in take-up. The gains are slightly larger for the multiplicative perturbation of the propensity score because it gives higher weight to people with high unobserved costs in our context. Hence, efforts to raise take-up would benefit immigrant children and hence, reduce the gaps in school performance between natives and immigrants. Yet, while these simulations are informative, they do not tell us which citizenship policies work and how take-up could be increased to reap the gains identified in Row (A). Would it be better to liberalize access to birthright citizenship for

[34] See Table 1 in Carneiro et al. (2011) for the corresponding weights to calculate MPRTE from MTE for small policy changes. Note that we do not consider marginal changes to our instruments as they are all discrete variables.

second-generation immigrants? Or should the government further facilitate the naturalization of first-generation immigrants and their offspring instead? We will now turn to marginal treatment response functions to answer these questions.

6.2 Using Marginal Treatment Response Functions to Separate the Gains from Birthright Citizenship and Naturalization

To inform policy-makers which citizenship policies work, we turn to the marginal treatment response approach recently developed by Mogstad et al. (2018, 2020). The idea is to obtain estimates of marginal treatment effects that are specific to a particular citizenship policy. In our case, we would like to separate the benefits of birthright citizenship from the gains of naturalization by first-generation immigrants. We start with splitting our instruments into two subsets: the first set describes pathways to naturalization for first-generation immigrants (individual and family eligibility); the second set describes access to citizenship for second-generation immigrants (birthright citizenship). The vector of instruments is now given by $\tilde{Z} = (Z_1, Z_2)$ where Z_1 contains the instruments for individual and family eligibility; and Z_2 denotes the indicator for birthright citizenship. Potential treatments are now defined as:

$$N_1(z_1) = \sum_{z_2} E[\mathbb{1}[Z_2 = z_2]N(z_1, Z_2, X)] \quad \text{and} \quad N_2(z_2) = \sum_{z_1} E[\mathbb{1}[Z_1 = z_1]N(z_2, Z_1, X)] \tag{6}$$

The sum in $N_1(z_1, x)$ is over $Z_2 = (0, 1)$, while the sum to obtain $N_2(z_2)$ is over the combinations of Z_2 containing the indicators for individual and family eligibility. The marginal potential treatments represent the take-up decision if the lth component of Z_l is set to z_l (where $l = 1, 2$), but all other components remain at their observed realizations. Using the partial monotonicity assumption, we can formulate the selection into treatment as:

$$N_l(z_l) = \mathbb{1}[U_l \leq p(z_l, Z_{-l}, X)] \tag{7}$$

where U_l is distributed uniformly on the unit interval conditional on any realization of Z_{-l}, X and $p(z_l, Z_{-l}, X)$ is the propensity score using the lth instrument. The unobservable U_l is now the unobserved costs of obtaining citizenship through margin l. In our case, U_1 denotes the unobserved cost of naturalization if an immigrant is eligible through individual or family eligibility. The observed take-up choice is characterized by (U_1, U_2), which measures the unobserved costs of citizenship obtained through naturalization or birthright, respectively.

The marginal treatment response functions (Mogstad et al., 2018, 2020) can then be defined as:

$$\begin{aligned} m_{l,0}(u_l, z_{-l}, x) &= E[Y_0 | U_l = u_l, Z_{-l} = z_{-l}, X = x] \quad \text{and} \\ m_{l,1}(u_l, z_{-l}, x) &= E[Y_1 | U_l = u_l, Z_{-l} = z_{-l}, X = x]. \end{aligned} \tag{8}$$

The MTR functions describe average outcomes in the treated and non-treated state as a function of the unobserved cost of obtaining citizenship along the lth margin (U_l) conditional on the other instruments (Z_{-l}) and control variables (X). From each pair of MTR functions we obtain a marginal treatment effect function ($m_{l,1} - m_{l,0}$) that is defined in terms of a different margin of selection (U_l) obtained from the lth instrument set.

After estimating the instrument-specific MTR functions, we simulate the effect of marginal policy changes that would raise the propensity score. The key difference to the simulations in the last section is that we calculate the gains from raising take-up using the instrument-specific MTE (rather than the MTE estimated from all instruments jointly). Hence, we consider an increase in the propensity score by α where α is either added to the propensity score or the propensity score is raised by a factor $(1 + \alpha)$. The results are shown in Row (B) of Table 6 for the MTE using birthright citizenship and Row (C) for the MTE using individual and family eligibility as instruments. As before, the returns to raising the propensity score multiplicatively are slightly larger because this perturbation puts more weight on children with high unobserved costs who have the highest potential benefits. Raising the take-up of citizenship through either birthright or naturalization benefits immigrant children. A comparison of Rows (B) and (C) reveals that the potential gains of take-up are larger for birthright citizenship than for naturalization.

One question that remains is then how policy-makers could achieve a higher take-up for the two citizenship policies. Citizenship take-up could increase by either expanding eligibility (e.g. by reducing the residency requirements necessary) or reducing the barriers to take-up (e.g. the need to renounce the original citizenship). With our data, we can shed light on the first channel: how liberalizing eligibility would influence take-up decisions. We consider two reforms of Germany's citizenship law that mimic current US policies. For birthright citizenship, we drop the 8-year residency requirements for parents of immigrant children born in the host country. The last Column of Row (B) shows that a US-style birthright citizenship would increase take-up by around 6 percentage points (from 0.833 in the sample to 0.891 under the alternative policy). For first-generation immigrants, we consider a reduction in the residency requirements from currently 8 years to 5 years. The last Column in Row (C) shows that this has no effect on the propensity score whatsoever indicating that residency requirements are not the binding constraint for the decision to naturalize in Germany.

Overall then, our policy simulations show that raising take-up would bring sizable gains in terms of improved educational performance. A comparison of birthright citizenship and naturalization for first-generation immigrants reveals that the potential benefits are larger for birthright citizenship than for naturalization for two reasons: because the improvements in school performance are more pronounced and because take-up can be more easily raised by changing eligibility criteria.

7 Conclusion

Granting access to citizenship is one of the most fundamental integration policies a host country can offer. In this paper, we shed light on the actual and potential gains from citizenship by estimating marginal returns (MTEs) and marginal treatment responses (MTRs) to a German passport. We study two important policy reforms in Germany, a country that has substantially liberalized its citizenship policy in recent decades.

We find strong evidence for a reverse selection on gains among immigrant children. Hence, children with high costs to treatment have higher returns than those with low resistance and a high likelihood of take-up. As a result, the average treatment effect on the untreated (ATU) is higher in absolute value than the average effect on the treated (ATT). The LATE is typically in the vicinity of the ATT and therefore misses the sizable potential gains for untreated children.

We find the strongest effects of treatment for weaker students through a decline in grade retention and an increase in the share attending high school. We also see positive returns through improvements in school grades on key subjects like Math. Interestingly, we see few effects of citizenship on more objective test scores. The pattern of improved school performance but few effects on analytical skills could indicate that teacher evaluations of immigrant children become more favorable if the child is naturalized. Another potential explanation is that citizenship affects the motivation of students, which could manifest in less disruptive behavior or more class participation.

Exploring the sources of unobservable costs to treatment indicates that the positive returns to citizenship are large for treated children from families with high resistance to citizenship. These are typically families with more favorable socio-economic characteristics, in which the mother is more educated and the family is more likely to speak German at home than families with low resistance. Making citizenship accessible to these families would reap substantial benefits for both the immigrants and the host society.

The fact that we observe reverse selection on unobservable gains raises the question whether a reform of citizenship policies would carry any additional benefits. Our policy simulations using marginal policy-relevant treatment effects yields two conclusions: first, an expansion of take-up would bring sizable gains in terms of improved school performance for immigrant children. Second, we find that the gains from expanding birthright citizenship are larger than for liberalizing naturalization among first-generation immigrants. Birthright citizenship not only promises sizable improvements in school performance; but it is also straightforward to increase take-up by dropping or reducing parental residency requirements.

References

Andresen, M. (2018). Exploring marginal treatment effects: Flexible estimation using Stata. *The Stata Journal*, 18(1):118–158.

Arnold, D., Dobbie, W., and Yang, C. S. (2018). Racial bias in bail decisions. *Quarterly Journal of Economics*, 133(4):1885–1932.

Autor, D., Kostl, A., Mogstad, M., and Setzler, B. (2019). Disability benefits, consumption insurance, and household labor supply. *American Economic Review*, 109(7):2613–54.

Avitabile, C., Clots-Figueras, I., and Masella, P. (2013). The effect of birthright citizenship on parental integration outcomes. *Journal of Law and Economics*, 56(3):777–810.

Avitabile, C., Clots-Figueras, I., and Masella, P. (2014). Citizenship, fertility, and parental investments. *American Economic Journal: Applied Economics*, 6(4):35–65.

Becker, G. S. and Lewis, H. G. (1973). On the interaction between the quantity and quality of children. *Journal of Political Economy*, 81(2):S279–S288.

Becker, G. S. and Tomes, N. (1976). Child endowments, and the quantity and quality of children. *Journal of Political Economy*, 84(4, Part 2):S143–S162.

Bhuller, M., Dahl, G. B., Løken, K. V., and Mogstad, M. (2020). Incarceration, recidivism, and employment. *Journal of Political Economy*, 128(4):1269–1324.

Bisin, A. and Verdier, T. (2011). The economics of cultural transmission and socialization. In Benhabib, J., Jackson, M., and Bisin, A., editors, *Handbook of Social Economics*, volume 1A, chapter 9, pages 339–416. Amsterdam: Elsevier.

Björklund, A. and Moffitt, R. (1987). The estimation of wage gains and welfare gains in self-selection models. *Review of Economics and Statistics*, 69(1):42–49.

Blossfeld, H.-P., Roßbach, H.-G., and von Maurice, J. E. (2011). Education as a Lifelong Process. The German National Educational Panel Study (NEPS). *Zeitschrift für Erziehungswissenschaft*, Special Issue 14, Suppl 2.

Borjas, G. J. (1992). Ethnic capital and intergenerational mobility. *Quarterly Journal of Economics*, 107(1):123–150.

Borjas, G. J. (1993). The intergenerational mobility of immigrants. *Journal of Labor Economics*, 11(1, Part 1):113–135.

Bratsberg, B., Ragan, Jr, J. F., and Nasir, Z. M. (2002). The effect of naturalization on wage growth: A panel study of young male immigrants. *Journal of Labor Economics*, 20(3):568–597.

Brinch, C. N., Mogstad, M., and Wiswall, M. (2017). Beyond LATE with a discrete instrument. *Journal of Political Economy*, 125(4):985–1039.

Carneiro, P., Heckman, J. J., and Vytlacil, E. J. (2011). Estimating marginal returns to education. *American Economic Review*, 101(6):2754–2781.

Carneiro, P. and Lee, S. (2009). Estimating distributions of potential outcomes using local instrumental variables with an application to changes in college enrollment and wage inequality. *Journal of Econometrics*, 149(2):191–208.

Carneiro, P., Lokshin, M., and Umapathi, N. (2017). Average and marginal returns to upper secondary schooling in Indonesia. *Journal of Applied Econometrics*, 32(1):16–36.

Chiswick, B. R. (1978). The effect of Americanization on the earnings of foreign-born men. *Journal of Political Economy*, 86(5):897–921.

Cornelissen, T., Dustmann, C., Raute, A., and Schönberg, U. (2018). Who benefits from universal child care? Estimating marginal returns to early child care attendance. *Journal of Political Economy*, 126(6):2356–2409.

Cunha, F. and Heckman, J. J. (2007). The technology of skill formation. *American Economic Review: Papers and Proceedings*, 97(2):31–47.

Dustmann, C. (2008). Return migration, investment in children, and intergenerational mobility. Comparing sons of foreign- and native-born fathers. *Journal of Human Resources*, 43(2):299–324.

Dustmann, C. and Glitz, A. (2011). Migration and education. In Hanushek, E., Machin, S., and Woessmann, L., editors, *Handbook of the Economics of Education*, volume 4, pages 327–439. Amsterdam: Elsevier.

Federal Office for Migration and Refugees (2019). *2016/2017 Migration Report: Key Results*. Bundesamt für Migration und Flüchtlinge (BAMF), Nürnberg.

Federal Statistical Office (2018). Bevölkerung und Erwerbstätigkeit. Einbürgerungen (Fachserie 1 Reihe 2.1).

Felfe, C. and Lalive, R. (2018). Does early child care help or hinder child development? *Journal of Public Economics*, 159:33–53.

Felfe, C., Rainer, H., and Saurer, J. (2020). Why birthright citizenship matters for immigrant children: Impacts on parental educational choices. *Journal of Labor Economics*, 38(1):143–182.

French, E. and Song, J. (2014). The effect of disability insurance receipt on labor supply. *American Economic Journal: Economic Policy*, 6(2):291–337.

Gathmann, C. and Keller, N. (2018). Access to citizenship and the economic assimilation of immigrants. *Economic Journal*, 128(616):3141–3181.

Gathmann, C., Keller, N., and Monscheuer, O. (2019). Citizenship and social integration. Mimeo, Heidelberg University.

Gathmann, C. and Monscheuer, O. (2021). Labor market assimilation of immigrants in Germany. Mimeo, Heidelberg University.

Hainmueller, J., Hangartner, D., and Pietrantuono, G. (2015). Naturalization fosters the long-term political integration of immigrants. *Proceedings of the National Academy of Sciences*, 112(41):12651–12656.

Hainmueller, J., Hangartner, D., and Pietrantuono, G. (2017). Catalyst or crown: Does naturalization promote the long-term social integration of immigrants? *American Political Science Review*, 111(2):256–276.

Heckman, J. J. (1997). Instrumental variables: A study of implicit behavioral assumptions used in making program evaluations. *Journal of Human Resources*, 32(3):441–462.

Heckman, J. J. and Vytlacil, E. J. (1999). Local instrumental variables and latent variable models for identifying and bounding treatment effects. *Proceedings of the National Academy of Sciences*, 96(8):4730–4734.

Heckman, J. J. and Vytlacil, E. J. (2001). Policy-relevant treatment effects. *American Economic Review: Papers and Proceedings*, 91(2):107–111.

Heckman, J. J. and Vytlacil, E. J. (2005). Structural equations, treatment effects and econometric policy evaluation. *Econometrica*, 73(3):669–738.

Heckman, J. J. and Vytlacil, E. J. (2007). Econometric evaluation of social programs, Part II: Using the marginal treatment effect to organize alternative econometric estimators to evaluate social programs, and to forecast their effects in new environments. In Heckman, J. J. and Leamer, E. E., editors, *Handbook of Econometrics*, volume 6, Part B, chapter 71, pages 4875–5143. Amsterdam: Elsevier.

Kamhöfer, D. A., Schmitz, H., and Westphal, M. (2018). Heterogeneity in marginal non-monetary returns to higher education. *Journal of the European Economic Association*, 17(1):205–244.

Kaufmann, K. M. (2014). Understanding the income gradient in college attendance in Mexico: The role of heterogeneity in expected returns. *Quantitative Economics*, 5(3):583–630.

Kline, P. and Walters, C. R. (2016). Evaluating public programs with close substitutes: The case of Head Start. *Quarterly Journal of Economics*, 131(4):1795–1848.

Lalonde, R. and Topel, R. (1997). Economic impact of international migration and the economic performance of migrants. In Rosenzweig, M. and Stark, O., editors, *Handbook of Population and Family Economics*, volume 1B, chapter 14. Amsterdam: Elsevier Science.

Maestas, N., Mullen, K. J., and Strand, A. (2013). Does disability insurance receipt discourage work? Using examiner assignment to estimate causal effects of SSDI receipt. *American Economic Review*, 103(5):1797–1829.

Moffitt, R. A. (2019). The marginal labor supply disincentives of welfare reforms. NBER Working Paper No. 26028, National Bureau of Economic Research.

Mogstad, M., Santos, A., and Torgovitsky, A. (2018). Using instrumental variables for inference about policy relevant treatment parameters. *Econometrica*, 86(5):1589–1619.

Mogstad, M. and Torgovitsky, A. (2018). Identification and extrapolation of causal effects with instrumental variables. *Annual Review of Economics*, 10:577–613.

Mogstad, M., Torgovitsky, A., and Walters, C. R. (2020). The causal interpretation of two-stage least squares with multiple instrumental variables. NBER Working Paper No. 25691, National Bureau of Economic Research.

Nybom, M. (2017). The distribution of lifetime earnings returns to college. *Journal of Labor Economics*, 35(4):903–952.

OECD (2018a). *International Migration Outlook 2018*. OECD Publishing: Paris, Paris.

OECD (2018b). *Settling In 2018: Indicators of Immigrant Integration*. OECD Publishing: Paris.

Shachar, A., Bauboeck, R., Bloemraad, I., and Vink, M., editors (2017). *The Oxford Handbook of Citizenship*. Oxford, UK: Oxford University Press.

Sweetman, A. and Van Ours, J. C. (2014). Immigration: What about the children and grandchildren? In Chiswick, B. and Miller, P. W., editors, *Handbook of the Economics of International Migration*, volume 1B, chapter 21. Elsevier, Amsterdam.

Torgovitsky, A. and Shea, J. (2019). IVMTE: A R package for implementing marginal treatment effect methods. doi 10.2139/ssrn.3516114.

Vytlacil, E. J. (2002). Independence, monotonicity, and latent index models: An equivalence result. *Econometrica*, 70(1):331–41.

Walters, C. R. (2018). The demand for effective charter schools. *Journal of Political Economy*, 126(6):2179–2223.

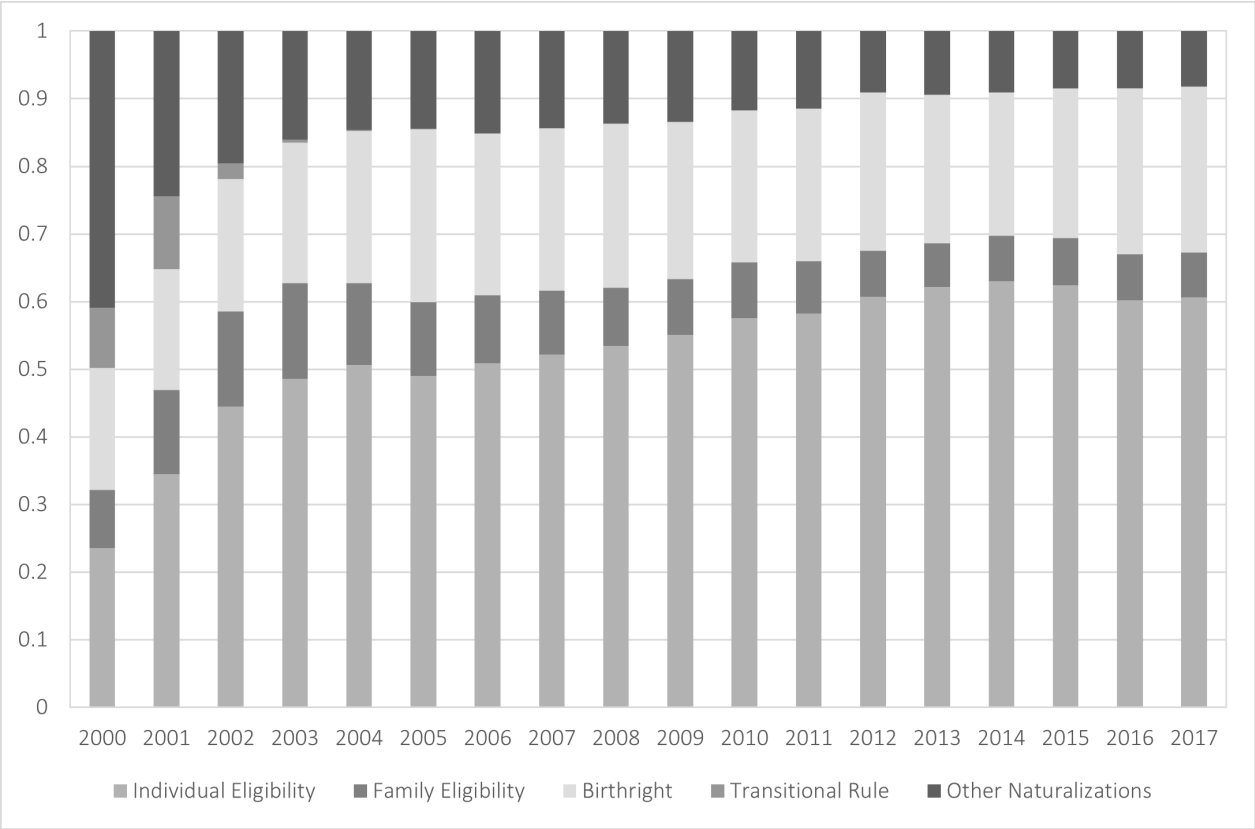


Figure 1: Naturalizations by Type of Eligibility

Notes: The figure shows the different pathways to citizenship in Germany: individual eligibility, as a family member of an eligible adult, by birthright citizenship or under the transitional rule of the 2000 reform. Data prior to 2000 are not available as they did not distinguish between different legal forms of citizenship acquisition.
Sources: Federal Office for Migration and Refugees (2019) and Federal Statistical Office (2018).

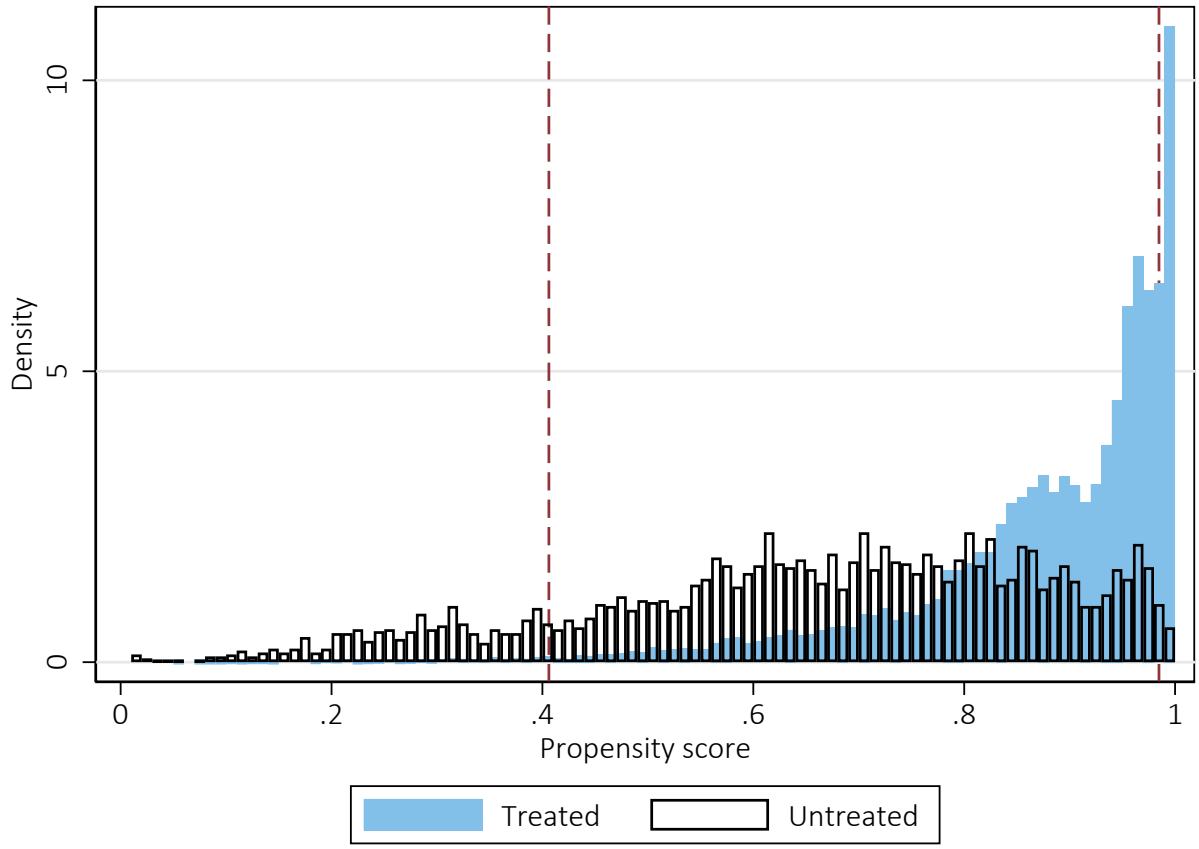
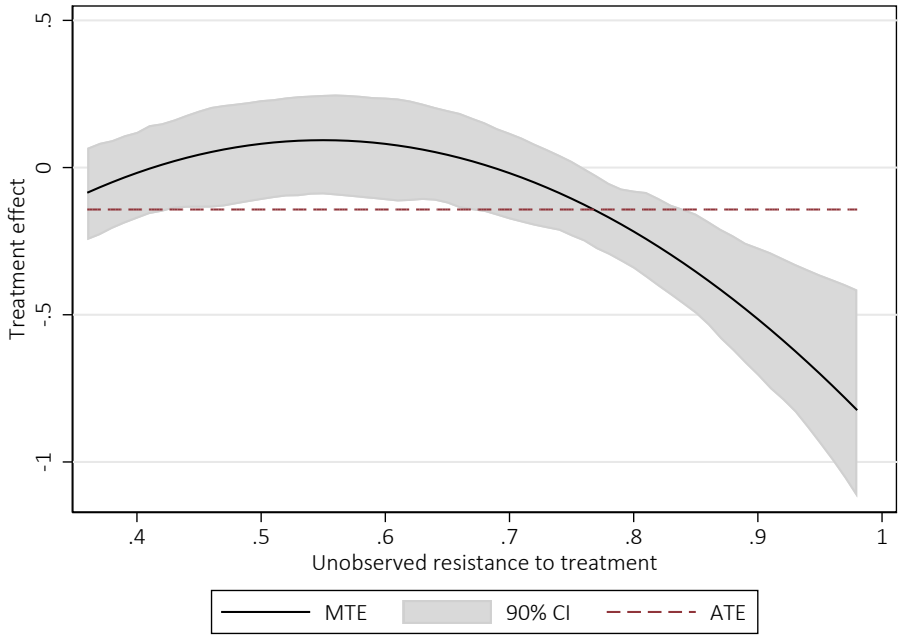
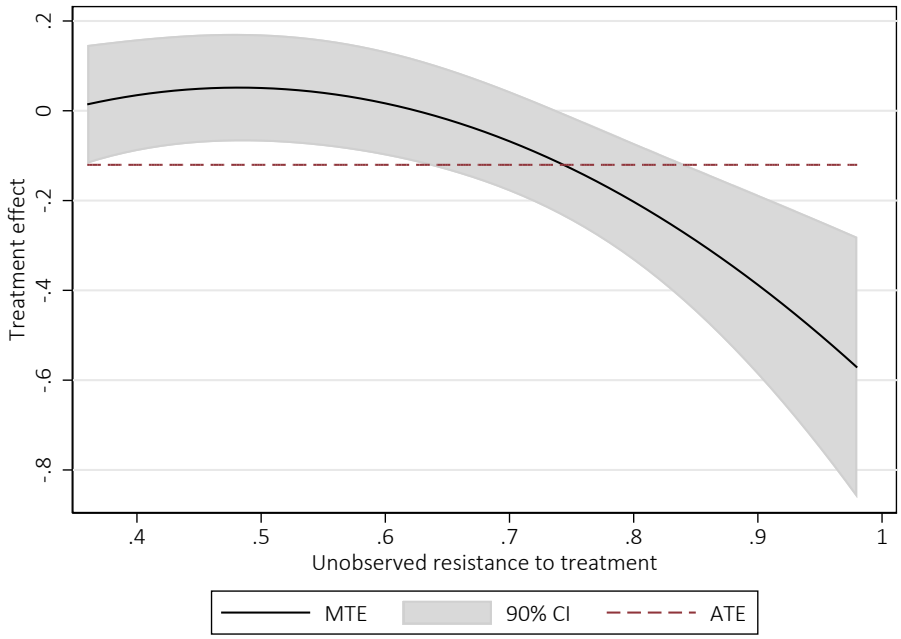


Figure 2: Common Support for Treated and Untreated Individuals

Notes: The figure shows the distribution of the propensity score for the full sample, separately for treated (solid gray bars) and non-treated (black-transparent bars) immigrant children.
Source: National Educational Panel Study, Starting Cohorts 2–4, 2010–2018.



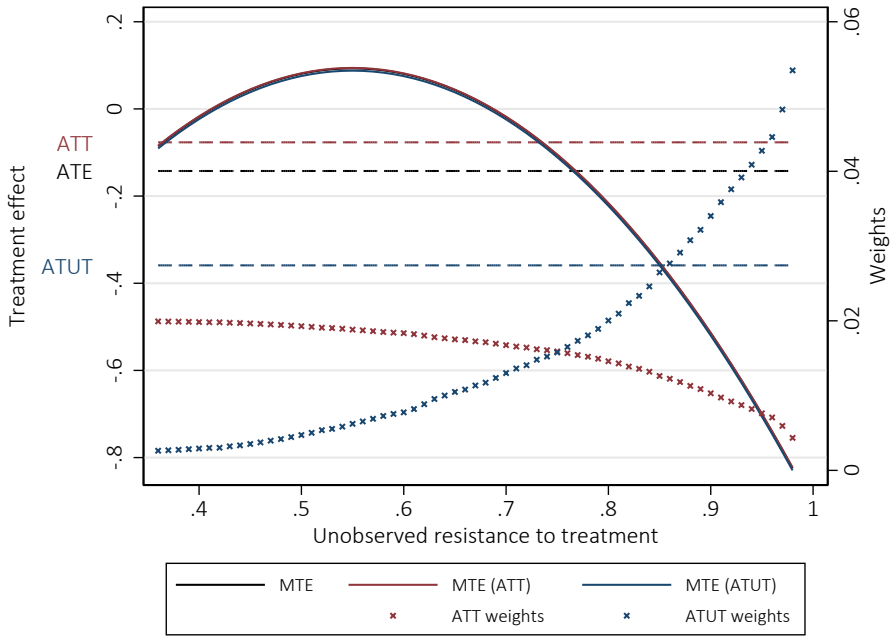
(a) MTE Using the Local Instrumental Variable Approach



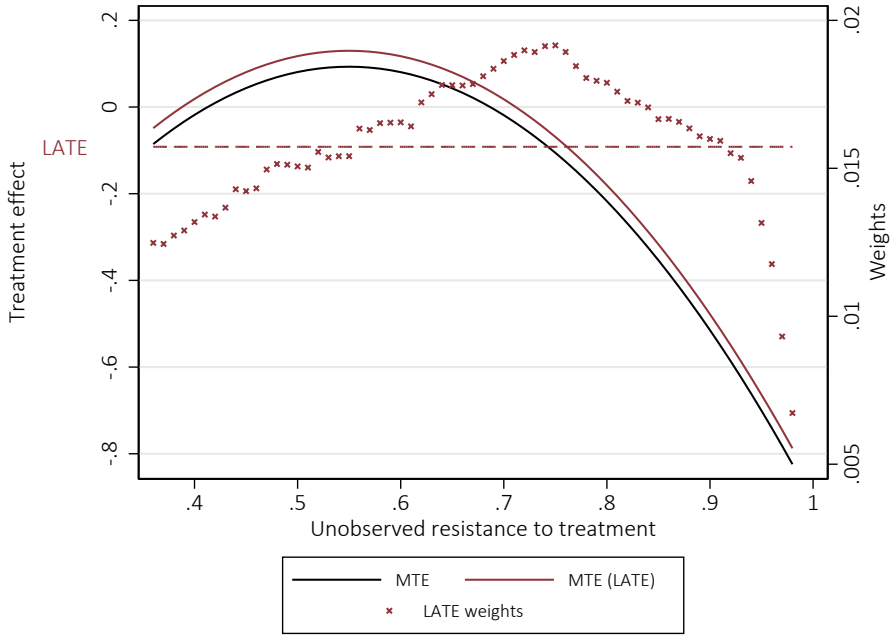
(b) MTE Using the Separate Estimation Approach

Figure 3: Marginal Treatment Effects for Grade Retention

Notes: The figure shows the marginal treatment effects for grade retention against the unobserved resistance (U_N) and the confidence intervals (shown in gray). The average treatment effect (ATE), shown as the red dashed line, is obtained by aggregating the MTE over the distribution of observable characteristics and unobservable resistance in the sample.
Source: National Educational Panel Study, Starting Cohorts 2–4, 2010–2018.



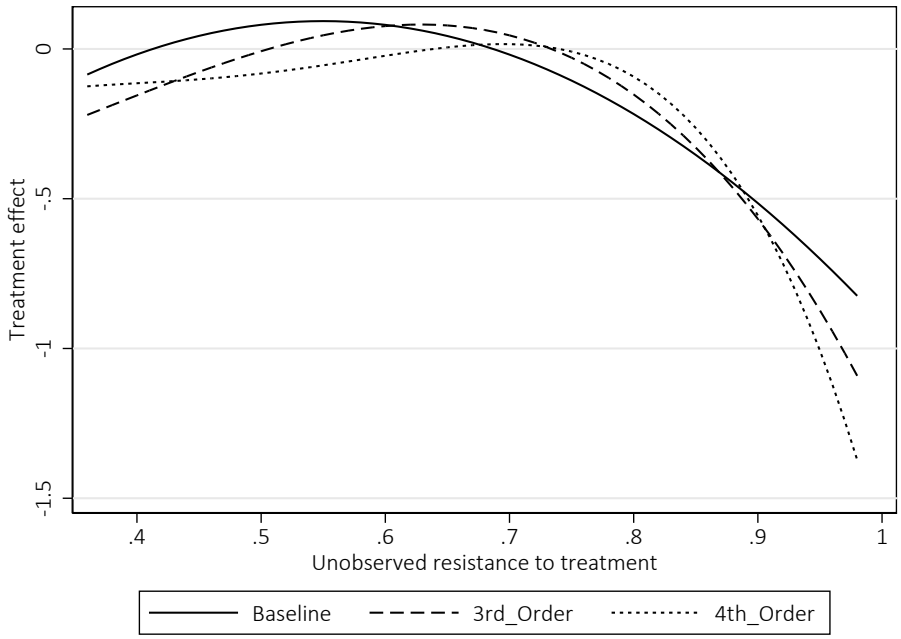
(a) ATT and ATU



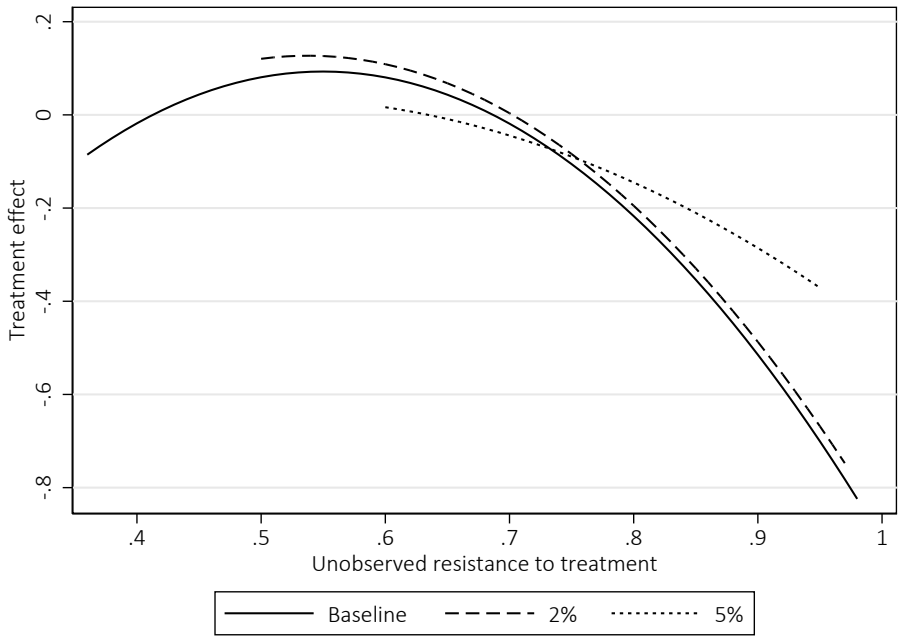
(b) LATE

Figure 4: Weights of Common Treatment Effects

Notes: The figure shows marginal treatment effects and the weights to calculate common treatment parameters from the MTEs against the unobserved resistance to treatment. Panel (a) shows the weights for ATT and ATU as well as the MTE if evaluated at the observable characteristics of the treated or untreated sample. Panel (b) shows the weights for the LATE and the MTE curve if evaluated with the observable characteristics of the complier population.
Source: National Educational Panel Study, Starting Cohorts 2–4, 2010–2018.



(a) Alternative Polynomials



(b) Alternative Trimming Margins

Figure 5: Specification Checks

Notes: Panel (a) compares the MTE curve for grade retention in the baseline where $k(P(Z) = K'(P(Z))$ is specified as a second-order polynomial with the MTE curves from specifying a third-order or fourth-order polynomials instead. Panel (b) compares the MTE curve for grade retention in the baseline where we impose a 1% trimming margin with the MTE curves generated using a 2% or 5% trimming margin instead.
Source: National Educational Panel Study, Starting Cohorts 2–4, 2010–2018.

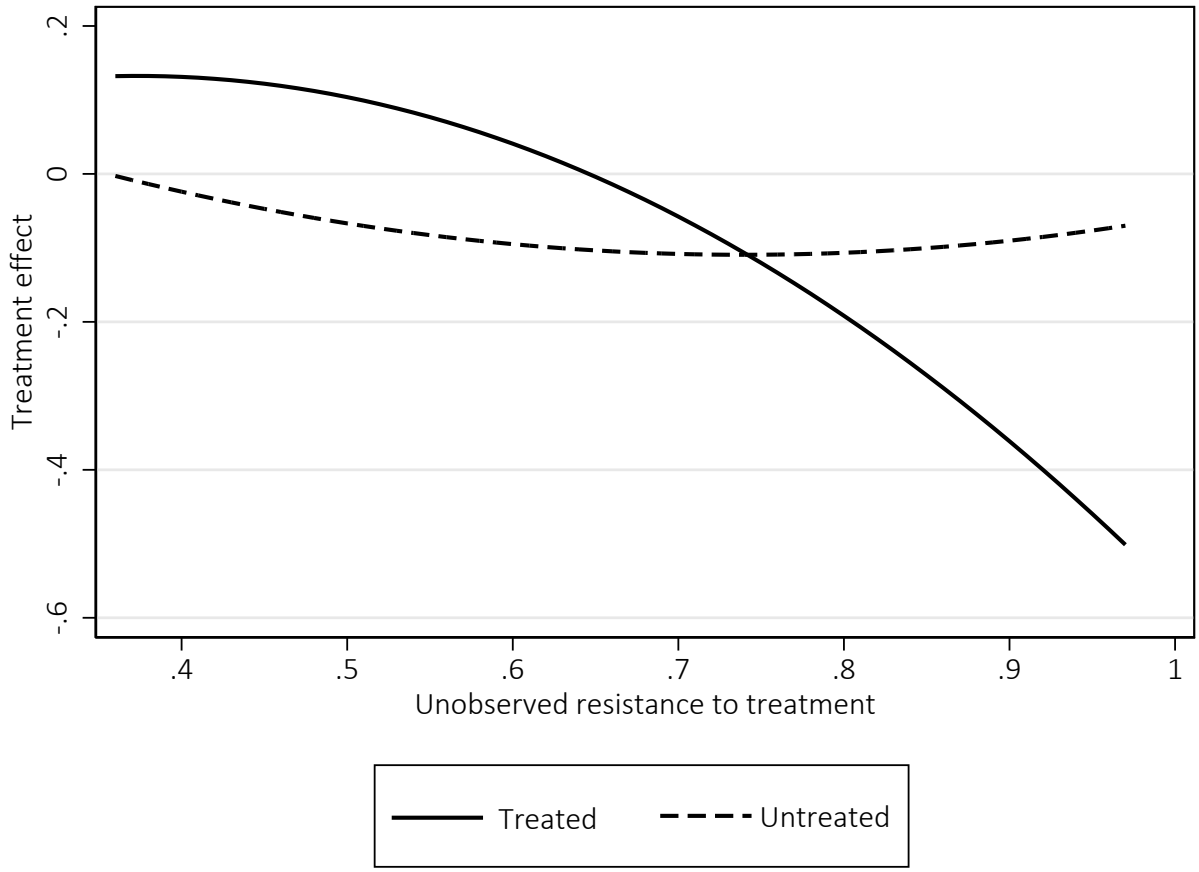
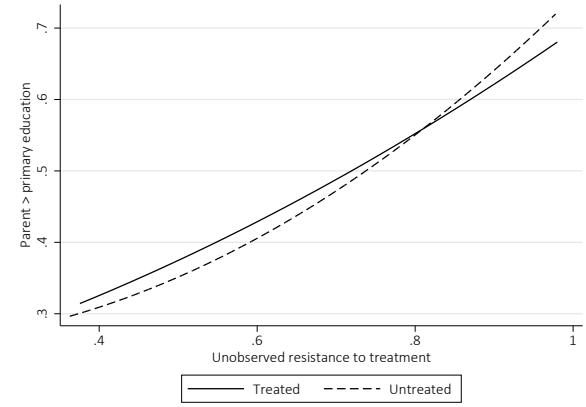
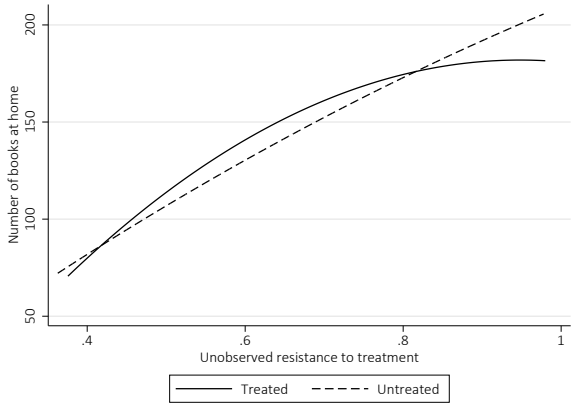


Figure 6: Unobserved Gains and Unobserved Resistance by Citizenship Status

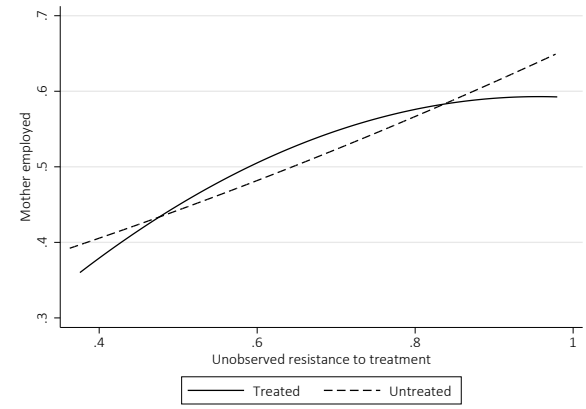
Notes: The figure plots the unobserved gains to treatment (on the y-axis) against the unobserved resistance to treatment (on the x-axis) separately for immigrant children with citizenship (the black line) and children without citizenship (the dashed line) for our main outcome grade retention. The graph is obtained from estimating the baseline model in equations (2) amd (5) using the separate approach.
Source: National Educational Panel Study, Starting Cohorts 2–4, 2010–2018.



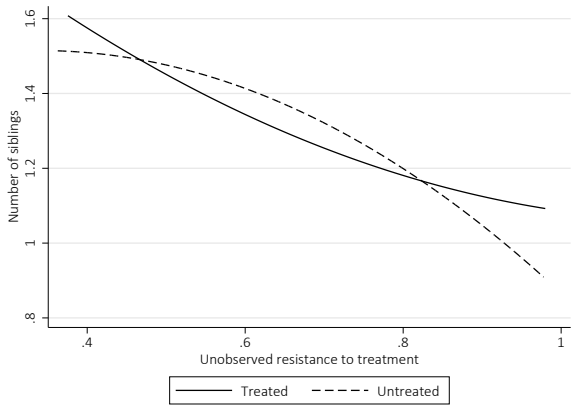
(a) Maternal Education



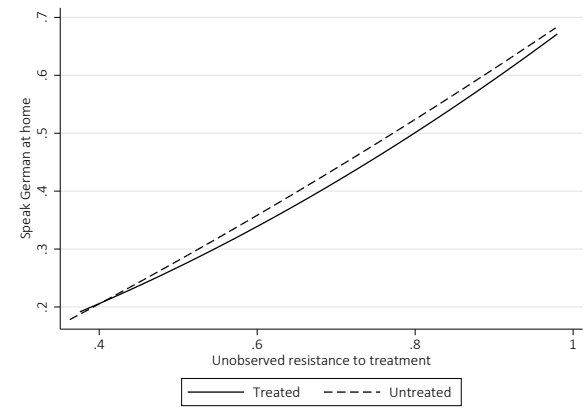
(b) Books at Home



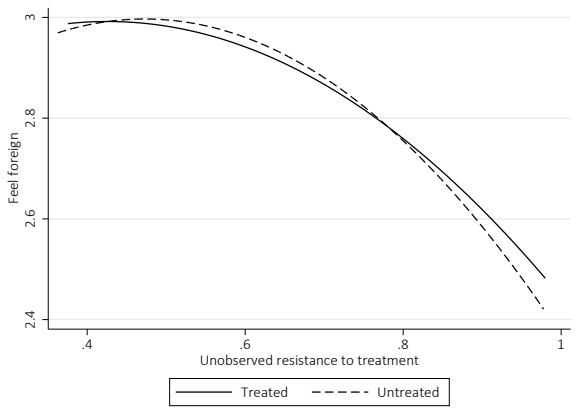
(c) Maternal Employment



(d) Number of Siblings



(e) Speak German at Home



(f) Feel Foreign in Germany

Figure 7: Family Characteristics against Unobserved Costs by Citizenship Status

Notes: The figure plots family characteristics (residualized from a regression on all control variables and instruments) on the y-axis against the unobserved barriers to citizenship, U_N, on the x-axis. Panel (a) shows whether the mother has more than primary education; Panel (b) the number of books at home. Panel (c) shows maternal employment and Panel (d) the number of siblings. Panel (e) shows whether the child speaks mostly German at home; Panel (f) shows whether the child feels as a foreigner in Germany on a scale from one to five wither higher values indicating a stronger feeling.
Source: National Educational Panel Study, Starting Cohorts 2–4, 2010–2018.

Table 1: Eligibility Rules for Citizenship

Children with Foreign-born Parents	**First- or Second-Generation Immigrant (Foreign- or German-born)**	**Second-Generation Immigrant (Born in Germany)**
1991 reform	**Individual Eligibility**	
	Child arrives at ages 0–7: At age 16	
	Child arrives at ages 8–14: After 8 years	
1991/2000 reform	**Family Eligibility**	
	Parent arrives at ages 8–14: After 8 years	
	Parent arrives at age 15 or older: After 15 (8) years prior to 2000 (since 2000)	
2000 reform		**Birthright Citizenship**
		Child born on or after January 1, 2000;
		Parent 8 years of legal residency

Notes: The table describes the pathways to citizenship for children with foreign-born parents after the 1991 and 2000 citizenship reforms. See Section 2 for a more detailed discussion of the reforms and eligibility rules for citizenship.

Table 2: Summary Statistics for Sample of Immigrant Children

	Mean	SD	Minimum	Maximum
Child outcomes				
Grade retention (ever repeated a grade)	0.16	0.36	0	1
Repeating current grade	0.02	0.12	0	1
Academic track	0.50	0.50	0	1
Academic track recommendation	0.52	0.50	0	1
Reading comprehension (standardized test)	−0.30	1.25	−4.7	5.5
Math (standardized test)	−0.21	1.16	−4.5	4.8
German grade	2.73	0.88	1	6
Math grade	2.77	1.05	1	6
Child treatment				
German citizenship	0.84	0.37	0	1
Child eligibility				
Birthright	0.34	0.47	0	1
Individual eligibility	0.34	0.47	0	1
Family eligibility	0.96	0.19	0	1
Child characteristics				
Girl	0.53	0.50	0	1
Foreign-born	0.15	0.35	0	1
Region of origin				
Traditional EU countries (EU–15)	0.12	0.33	0	1
Turkey	0.17	0.38	0	1
Former Soviet Union	0.26	0.44	0	1
Central and Eastern Europe	0.21	0.41	0	1
Other countries	0.23	0.42	0	1
Age at immigration if foreign-born	4.48	3.60	0	16
Birth cohort	2000	4.49	1993	2007
Age	13.21	4.01	4.3	24
Parental characteristics				
German citizenship	0.54	0.50	0	1
Years since immigration	21.82	9.69	0	58
Immigration year	1991	9.74	1957	2013
Birth cohort	1971	6.40	1949	1996
Age	41.63	6.20	14	65
Parental education				
Low	0.39	0.49	0	1
Medium	0.54	0.50	0	1
High	0.07	0.26	0	1
Observations	18,707			

Notes: The sample is restricted to first- and second-generation immigrant children with at least one foreign-born parent. The eligibility variables are binary indicators equal to one if the child is eligible under one of the access options specified (see Section 2 in the main text). Parents have low education if they have no secondary school degree or vocational degree, medium if they have a high school or vocational degree and high if they have a university or college degree.

Source: National Educational Panel Study, Starting Cohorts 2–4, 2010–2018.

Table 3: Selection Equation

	(1) Full Sample	(2) Full Sample	(3) Grade Retention Sub-Sample	(4) Grade Retention Sub-Sample
Child eligibility				
Birthright	0.134***	0.126***	0.119***	0.124***
	(0.009)	(0.008)	(0.022)	(0.021)
Individual eligibility	−0.028**	−0.022*	−0.266***	−0.266***
	(0.013)	(0.012)	(0.067)	(0.064)
Individual*Foreign-born	0.117***	0.065***	0.344***	0.222***
	(0.013)	(0.012)	(0.039)	(0.037)
Family eligibility	0.150***	0.057***	0.048	0.034
	(0.015)	(0.014)	(0.047)	(0.046)
Child characteristics				
Girl	0.010*	0.006	0.009	−0.002
	(0.005)	(0.005)	(0.012)	(0.011)
Foreign-born	−0.145***	−0.159***	−0.378***	−0.309***
	(0.011)	(0.011)	(0.037)	(0.034)
Region of origin (reference: Other countries)				
EU countries	0.037***		0.040***	
	(0.005)		(0.013)	
Traditional EU countries (EU–15)		−0.076***		−0.088***
		(0.008)		(0.020)
Central and Eastern Europe		−0.038***		−0.044**
		(0.007)		(0.017)
Former Soviet Union		0.149***		0.149***
		(0.008)		(0.020)
Turkey		−0.165***		−0.170***
		(0.007)		(0.017)
Age	−0.144***	−0.128***	−0.462*	−0.117
	(0.026)	(0.025)	(0.262)	(0.218)
Age squared	0.003***	0.002***	0.011	0.002
	(0.001)	(0.001)	(0.008)	(0.007)
Age at immigration	−0.005***	−0.003**	−0.003	−0.008
	(0.002)	(0.002)	(0.005)	(0.005)
Parental characteristics				
Years since immigration	0.005***	0.012***	0.008**	0.011***
	(0.001)	(0.001)	(0.004)	(0.004)
Years since immigration squared	−0.000***	−0.000***	−0.000**	−0.000**
	(0.000)	(0.000)	(0.000)	(0.000)
Child's year of birth FE	Yes	Yes	Yes	Yes
Survey cohort*year FE	Yes	Yes	No	No
Survey cohort FE	No	No	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Observations	18,479	18,480	3,219	3,219
Lower limit of common support	0.43	0.41	0.40	0.36
Upper limit of common support	0.97	0.98	0.98	0.98
Chi^2-Statistic for Z	519.212	328.783	150.311	101.051
P-value Chi^2	0.000	0.000	0.000	0.000

Notes: The table reports average marginal effects of a probit selection model where the dependent variable is whether the child has German citizenship or not. Columns (1) and (2) use the full sample including all survey cohorts and waves, while Columns (3) and (4) use sub-sample with valid values for out main outcome grade retention. The instruments are three binary indicators equal to one if a child is eligible for citizenship through a specific channel; and zero otherwise. The channels are: birthright citizenship (child born in the country on January 1, 2000 or later with one foreign-born parent who has lived in Germany legally for at least eight years); individual eligibility (child is at least 16 and has lived legally in the country for at least eight years); and family eligibility (child is under 18 and has one foreign-born parent who has lived in Germany legally for between eight and sixteen years (see Table 1 for details). We further interact individual eligibility with an indicator whether the child is born abroad to allow for differential take-up rates among foreign- and German-born children. Age at immigration of the child is equal to zero for children born in Germany. All specifications further include birth cohort fixed effects, year*survey cohort fixed effects, and state fixed effects. Bootstrapped standard errors are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.
Source: National Educational Panel Study, Starting Cohorts 2–4, 2010–2018.

Table 4: MTE Estimates for Main Outcomes

	(1) Grade Retention	(2) Grade Retention
Untreated State (β_0)		
Girl	0.010	−0.040
	(0.062)	(0.053)
Country of origin (reference: Other countries)		
EU countries	−0.083	
	(0.066)	
Traditional EU countries (EU–15)		0.033
		(0.094)
Central and Eastern Europe		−0.146*
		(0.088)
Former Soviet Union		0.146
		(0.118)
Turkey		0.095
		(0.109)
Treated State ($\beta_1 - \beta_0$)		
Girl*Propensity score	−0.010	0.045
	(0.068)	(0.059)
Country of origin (reference: Other countries)		
EU*Propensity score	0.094	
	(0.074)	
Traditional EU countries*Propensity score		−0.087
		(0.103)
Central and Eastern Europe*Propensity score		0.130
		(0.097)
Former Soviet Union*Propensity score		−0.121
		(0.124)
Turkey*Propensity score		−0.151
		(0.132)
Propensity score	4.588*	5.448***
	(2.590)	(2.029)
Propensity score squared	−4.886**	−4.955***
	(2.169)	(1.727)
Child's age and age of immigration	Yes	Yes
Parental years since migration	Yes	Yes
Child's year of birth FE	Yes	Yes
Survey cohort FE	Yes	Yes
State FE	Yes	Yes
Observations	3,219	3,219
P-value observable heterogeneity	0.411	0.049
P-value essential heterogeneity	0.000	0.009

Notes: The table reports estimates from the outcome equation where the dependent variable is grade retention, i.e., an indicator whether a child has ever repeated a grade. Coefficients not interacted with the propensity score measure the effect in the untreated state, while coefficients interacted with the propensity score show the differential effect between treatment (host country citizenship) and no treatment. All specifications also include child age and age squared, the child's age at immigration, birth cohort fixed effects, parental years since migration (linear and squared terms), year*survey cohort and federal state fixed effects. All additional control variables are restricted to have the same effect in the treated and untreated state. Bootstrapped standard errors are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Source: National Educational Panel Study, Starting Cohorts 2–4, 2010–2018.

Table 5: Treatment Effects for School Performance and Test Scores

	(1) Grade Retention	(2) Academic Track	(3) German Grade	(4) Reading Test	(5) Math Grade	(6) Math Test
ATE	−0.142**	−0.001	0.035	−0.200	−0.219*	−0.120
	(0.063)	(0.054)	(0.103)	(0.269)	(0.130)	(0.162)
ATT	−0.077	−0.061	0.095	−0.176	−0.179	−0.214
	(0.069)	(0.055)	(0.106)	(0.272)	(0.133)	(0.163)
ATU	−0.359***	0.154**	−0.139	−0.266	−0.324*	0.204
	(0.094)	(0.070)	(0.139)	(0.325)	(0.176)	(0.249)
LATE	−0.092	0.080	−0.069	−0.169	−0.313**	0.008
	(0.061)	(0.058)	(0.112)	(0.258)	(0.140)	(0.164)
Observations	3,219	13,887	11,613	4,534	11,625	7,316
P-value observable heterogeneity	0.049	0.000	0.000	0.094	0.000	0.003
P-value essential heterogeneity	0.009	0.000	0.113	0.817	0.026	0.121

Notes: The table reports estimates from the outcome equation in equation 3 where the dependent variables are grade retention (Column (1)), whether a child attends the academic track, i.e., high school (Column (2)); the grade in German (Column (3)) and in Math (Column (4)); the test scores in reading comprehension (Column (5)) and in Math skills (Column (6)). The ATE, ATT and ATU are weighted averages of the MTE estimates over both observed and unobserved characteristics for all children in the sample (for the ATE), for the children with German citizenship (for the ATT) and for children without German citizenship (for the ATU). The LATE is calculated by aggregating the MTE over the range of observed and unobserved characteristics of compliers, i.e., those individuals that switch treatment status when one of our instruments is turned on. In addition, we report the p-value of the null hypotheses of no observable or no essential (unobserved) heterogeneity. The specification is the same as in Table 4. Bootstrapped standard errors are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.
Source: National Educational Panel Study, Starting Cohorts 2–4, 2010–2018.

Table 6: Simulations of Alternative Citizenship Policies

	Marginal Policy-Relevant Treatment Effects		Propensity Score
	$P^{\alpha} = P + \alpha$	$P^{\alpha} = (1 + \alpha) * P$	(Sample: 0.833)
(A) Increase Citizenship Take-up	-0.057^{***}	-0.064^{***}	
	(0.015)	(0.017)	
(B) Liberalize Birthright Citizenship	-0.044^{**}	-0.050^{**}	
	(0.020)	(0.021)	
US-style Birthright Citizenship			0.891
(C) Facilitate Naturalizations	-0.032^{*}	-0.036^{*}	
	(0.018)	(0.020)	
Reduce Residency Requirements			0.833

Notes: The table shows simulations of alternative citizenship policies. The first columns report marginal policy-relevant treatment effects (MPRTE) of increasing citizenship take-up on the probability of grade retention from alternative citizenship policies. Take-up is raised either by adding α to the propensity score or multiplying the propensity score by $(1 + \alpha)$ where $\alpha \to 0$. The first policy experiment (in Row (A)) generates the MPRTE from weighting the MTE estimated from using all instruments simultaneously. The second policy experiment (in Row (B)) estimates the MRPTE from the marginal treatment response function using birthright citizenship as the instrument while controlling for individual and family eligibility. The third policy experiment (in Row (C)) estimates the MPRTE from the marginal treatment response function using individual and family eligibility as the instrument while controlling for birthright citizenship. The last column then reports how alternative citizenship reforms could affect the take-up decision. Row (B) implements a birthright citizenship policy as in the US where all immigrant children born in the host country receive citizenship (hence, dropping the 8-year residency requirements for the parent), Row (C) shows the take-up if we reduce the residency requirement for naturalizations from 8 years to 5 years. Bootstrapped standard errors are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.
Source: National Educational Panel Study, Starting Cohorts 2–4, 2010–2018.

A Appendix

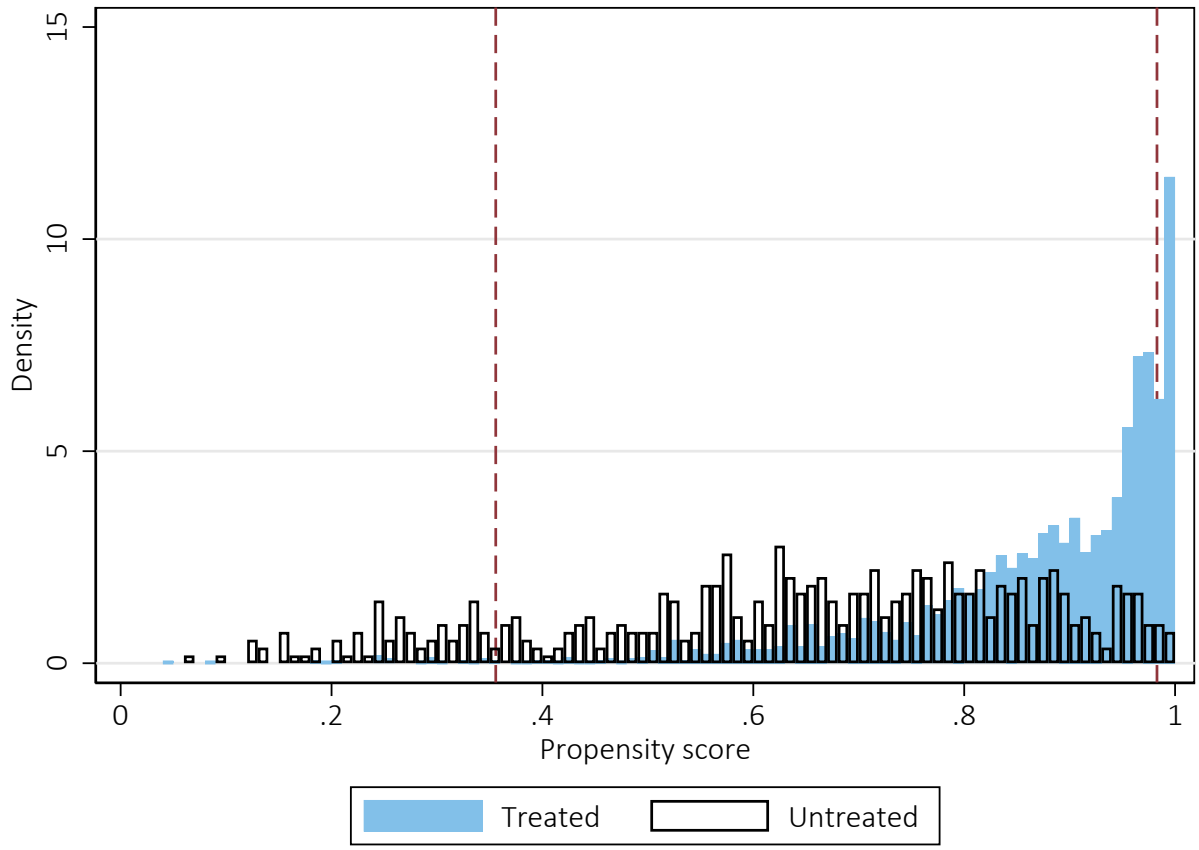


Figure A1: Common Support for Grade Retention Sub-sample

Notes: The figure shows the distribution of the propensity score for the sub-sample with valid information on grade retention, separately for treated (solid gray bars) and non-treated (black-transparent bars) immigrant children.

Source: National Educational Panel Study, Starting Cohorts 2–4, 2010–2018.

Table A1: Monotonicity

	Overall Sample				Grade Retention Sub-Sample			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Child eligibility								
Birthright	0.090***	0.099***			0.078***	0.109***		
	(0.008)	(0.008)			(0.020)	(0.020)		
Individual eligibility	−0.033**		−0.046***		−0.261***		−0.270***	
	(0.014)		(0.014)		(0.067)		(0.067)	
Individual*Foreign-born	0.125***		0.169***		0.315***		0.354***	
	(0.015)		(0.015)		(0.041)		(0.040)	
Family eligibility	0.121***			0.146***	0.064			0.098*
	(0.017)			(0.017)	(0.054)			(0.054)
Child characteristics								
Girl	0.006	0.005	0.005	0.006	0.000	−0.002	0.000	−0.001
	(0.005)	(0.005)	(0.005)	(0.005)	(0.012)	(0.012)	(0.012)	(0.012)
Foreign-born	−0.222***	−0.157***	−0.247***	−0.181***	−0.406***	−0.176***	−0.442***	−0.203***
	(0.013)	(0.011)	(0.013)	(0.011)	(0.041)	(0.027)	(0.040)	(0.027)
Region of origin (reference: Other countries)								
Traditional EU countries (EU–15)	−0.084***	−0.088***	−0.093***	−0.093***	−0.090***	−0.097***	−0.094***	−0.104***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.022)	(0.022)	(0.022)	(0.022)
Central and Eastern Europe	−0.037***	−0.040***	−0.036***	−0.035***	−0.043**	−0.047***	−0.040**	−0.044**
	(0.007)	(0.007)	(0.007)	(0.007)	(0.018)	(0.018)	(0.018)	(0.018)
Former Soviet Union	0.092***	0.097***	0.092***	0.097***	0.085***	0.098***	0.085***	0.100***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.017)	(0.017)	(0.017)	(0.017)
Turkey	−0.195***	−0.197***	−0.196***	−0.196***	−0.200***	−0.203***	−0.200***	−0.203***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.019)	(0.019)	(0.019)	(0.019)
Age	−0.097***	−0.121***	−0.118***	−0.138***	0.047	−0.014	0.077	0.021
	(0.025)	(0.025)	(0.025)	(0.025)	(0.196)	(0.198)	(0.196)	(0.199)
Age squared	0.002**	0.003***	0.002*	0.002***	−0.003	−0.001	−0.005	−0.003
	(0.001)	(0.001)	(0.001)	(0.001)	(0.006)	(0.006)	(0.006)	(0.006)
Age at immigration	−0.006***	−0.011***	−0.005***	0.000	−0.009	−0.007	−0.006	0.003
	(0.002)	(0.002)	(0.002)	(0.002)	(0.006)	(0.005)	(0.006)	(0.005)
Parental characteristics								
Years since immigration	0.007***	0.012***	0.016***	0.014***	0.008*	0.011***	0.013***	0.016***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.004)	(0.004)	(0.004)	(0.004)
Years since immigration squared	−0.000***	−0.000***	−0.000***	−0.000***	−0.000	−0.000**	−0.000**	−0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Child's year of birth FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Survey cohort*year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	18,707	18,707	18,707	18,707	3,260	3,260	3,260	3,260
R^2	0.193	0.186	0.185	0.183	0.203	0.188	0.199	0.180

Notes: The table shows estimates from a regression of actual treatment status (=1 if a person has German citizenship) on the three binary eligibility variables. For individual eligibility, an interaction with the indicator for being foreign-born is included. Columns (1) to (4) refer to the overall sample, Columns (5) to (8) refer to the sub-sample with information on grade retention. *** p<0.01, ** p<0.05, * p<0.1.

Source: National Educational Panel Study, Starting Cohorts 2–4, 2010–2018.

Table A2: Sample Attrition

	(1) Temporary Drop-Out	(2) Permanent Drop-Out
Child eligibility		
Birthright	0.046**	−0.019
	(0.023)	(0.014)
Individual eligibility	−0.017	−0.018
	(0.098)	(0.057)
Individual*Foreign-born	0.056	−0.032
	(0.047)	(0.028)
Family eligibility	0.063	−0.032
	(0.070)	(0.040)
Child characteristics		
Girl	0.016	0.008
	(0.015)	(0.009)
Foreign-born	−0.012	0.020
	(0.044)	(0.026)
Country of origin (reference: Other countries)		
Traditional EU countries (EU–15)	0.033	0.007
	(0.027)	(0.016)
Central and Eastern Europe	−0.017	0.011
	(0.022)	(0.013)
Former Soviet Union	−0.046**	0.016
	(0.021)	(0.012)
Turkey	0.065***	0.007
	(0.024)	(0.014)
Age	1.163***	0.000
	(0.242)	(0.141)
Age squared	−0.027***	0.003
	(0.008)	(0.004)
Age at immigration	−0.004	−0.008
	(0.008)	(0.005)
Parental characteristics		
Years since immigration	−0.009*	0.001
	(0.005)	(0.003)
Years since immigration squared	0.000	−0.000
	(0.000)	(0.000)
Child's year of birth FE	Yes	Yes
Survey cohort FE	Yes	Yes
State FE	Yes	Yes
Adjusted R^2	0.31	0.06
Mean dependent variable	0.36	0.06
Observations	2,943	2,943

Notes: The table shows estimates from a linear probability model where the dependent variable is equal to one if a child misses a wave (in Column (1)) or if a child drops out of the sample permanently (in Column (2)). The number of observations corresponds to the number of children in the sample. *** p<0.01, ** p<0.05, * p<0.1.

Source: National Educational Panel Study, Starting Cohorts 2–4, 2010–2018.

Table A3: Estimates for Other Educational Outcomes

	(1) Academic Track	(2) German Grade	(3) Reading Test	(4) Math Grade	(5) Math Test
Untreated State (β_0)					
Girl	0.050	−0.402***	0.269*	0.116	−0.173
	(0.033)	(0.071)	(0.162)	(0.088)	(0.119)
Country of origin (reference: Other countries)					
Traditional EU countries (EU–15)	−0.140**	0.162	−0.212	0.149	−0.267
	(0.057)	(0.129)	(0.297)	(0.174)	(0.207)
Central and Eastern Europe	−0.336***	0.407***	−0.882***	0.227	−0.614***
	(0.059)	(0.126)	(0.313)	(0.158)	(0.186)
Former Soviet Union	−0.197**	−0.016	−0.394	−0.041	−0.218
	(0.081)	(0.149)	(0.383)	(0.206)	(0.226)
Turkey	−0.289***	0.175	−0.897***	−0.191	−0.855***
	(0.057)	(0.129)	(0.272)	(0.159)	(0.193)
Treated State ($\beta_1 - \beta_0$)					
Girl*Propensity score	−0.039	0.143*	−0.168	0.093	−0.211
	(0.039)	(0.083)	(0.188)	(0.101)	(0.139)
Country of origin (reference: Other countries)					
Traditional EU countries*Propensity score	0.207***	−0.264*	0.504	−0.247	0.606**
	(0.066)	(0.151)	(0.351)	(0.205)	(0.252)
Central and Eastern Europe*Propensity score	0.388***	−0.554***	0.983***	−0.386**	0.869***
	(0.067)	(0.142)	(0.361)	(0.179)	(0.214)
Former Soviet Union*Propensity score	0.155*	0.046	0.443	0.084	0.332
	(0.085)	(0.157)	(0.397)	(0.222)	(0.239)
Turkey*Propensity score	0.195***	0.074	0.176	0.336*	0.310
	(0.071)	(0.155)	(0.323)	(0.192)	(0.229)
Propensity score	−3.919***	2.760	3.634	−2.438	−2.658
	(1.105)	(2.565)	(5.782)	(3.490)	(4.652)
Propensity score squared	3.552***	−2.896	−2.823	0.884	3.367
	(0.933)	(2.113)	(4.659)	(2.849)	(3.887)
Child's age and age of immigration	Yes	Yes	Yes	Yes	Yes
Parental years since migration	Yes	Yes	Yes	Yes	Yes
Child's year of birth FE	Yes	Yes	Yes	Yes	Yes
Survey cohort FE	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes
Observations	13,887	11,613	4,534	11,625	7,316

Notes: The table reports estimates from the outcome equation where the dependent variable is specified in the column header. Coefficients not interacted with the propensity score measure the effect in the untreated state, while coefficients interacted with the propensity score show the differential effect between treatment (host country citizenship) and no treatment. All specifications also include child age and age squared, the child's age at immigration, birth cohort fixed effects, parental years since migration (linear and squared terms), year*survey cohort and federal state fixed effects. All additional control variables are restricted to have the same effect in the treated and untreated state. Bootstrapped standard errors are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Source: National Educational Panel Study, Starting Cohorts 2–4, 2010–2018.

Table A4: Robustness Checks

	(1) Semiparam. Approach	(2) Separate Approach	(3) Track Control	(4) Birth Month Dummies	(5) Drop 12/99 and 01/00	(6) Drop Born After 04/00	(7) Drop Early Arrivals	(8) Drop Upper Secondary	(9) Drop Adult Children	(10) Drop Parent Info Missing	(11) Drop Ethnic Germans
ATE	−0.480***	−0.117**	−0.125*	−0.180***	−0.141**	−0.171	−0.165***	−0.142**	−0.135**	−0.134**	−0.217***
	(0.052)	(0.057)	(0.066)	(0.056)	(0.065)	(0.146)	(0.062)	(0.066)	(0.067)	(0.061)	(0.065)
ATT	−0.379***	−0.067	−0.055	−0.097	−0.081	−0.057	−0.096	−0.077	−0.076	−0.050	−0.128*
	(0.053)	(0.056)	(0.070)	(0.062)	(0.072)	(0.153)	(0.067)	(0.069)	(0.069)	(0.066)	(0.067)
ATU	−0.817***	−0.288***	−0.363***	−0.464***	−0.344***	−0.438**	−0.409***	−0.359***	−0.334***	−0.401***	−0.435***
	(0.086)	(0.090)	(0.096)	(0.093)	(0.099)	(0.188)	(0.108)	(0.104)	(0.114)	(0.108)	(0.108)
LATE	−0.453***	−0.117**	−0.065	−0.148**	−0.099	−0.237	−0.104*	−0.092	−0.069	−0.071	−0.188***
	(0.048)	(0.058)	(0.065)	(0.059)	(0.065)	(0.146)	(0.060)	(0.064)	(0.068)	(0.064)	(0.067)
Observations	3,219	3,219	3,040	3,220	3,200	1,754	2,951	3,219	3,006	2,922	2,329
P-value obs. het.	0.052	0.002	0.011	0.010	0.036	0.221	0.066	0.069	0.062	0.070	0.004
P-value ess. het.	0.000	0.016	0.001	0.001	0.022	0.054	0.011	0.008	0.036	0.007	0.009

Notes: The table reports estimates from the outcome equation where the specification is shown in the top row. Column (1) uses a semiparametric approach based on the double residual method. Column (2) estimates the relationship for treated and untreated state separately and calculates the MTE as their difference. Column (3) adds a binary track control variable (academic vs. non-academic track). Column (4) includes birth month dummies in the specification, while Column (5) drops children born around the cutoff date for birthright citizenship ("donut"). Column (6) drops all children born more than four months after the introduction of birthright citizenship. Column (7) drops children whose parents immigrated to Germany 1975 or before, while Column (8) drops all observations after the end of compulsory schooling. Column (9) drops children older than 18 years of age from the sample. Column (10) drops children with missing info on the country of birth for the second parent and Column (11) drops all families who might have entered as ethnic Germans. The ATE, ATT and ATU are weighted averages of the MTE estimates over both observed and unobserved characteristics for all children in the sample (for the ATE), for the children with German citizenship (for the ATT) and for children without German citizenship (for the ATU). The LATE is calculated by aggregating the MTE over the range of observed and unobserved characteristics of compliers, i.e., those individuals that are induced to change from no treatment to treatment when the instruments are turned on. In addition, we report the p-value of the null hypotheses of no observable or no essential (unobserved) heterogeneity. Bootstrapped standard errors are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Source: National Educational Panel Study, Starting Cohorts 2–4, 2010–2018.