ON LINE APPENDIX: Citizenship, Fertility and Parental Investments

By Ciro Avitabile, Irma Clots-Figueras and Paolo Masella

Appendix A. Alternative Comparison Group

The 2000 reform changed the citizenship status of children born to households, where at least one of the partners had been in Germany for 8 or more years. Couples where none of the partners has lived in Germany for 8 years were not directly targeted by the reform and may represent an alternative comparison group. Immigrant couples where none of the partners has been in Germany for 8 or more years are likely to show similarities with the treatment group in terms of current economic conditions, future economic opportunities, educational and cultural background and therefore can help to assess to what extent our main results are the artifact of economic and policy shocks that affected the immigrant population differentially.

However, there are at least three reasons why we expect the fertility behavior of immigrant couples with less than 8 years in Germany to respond to the citizenship reform. First, the composition of the immigrant population in terms of child preferences might change as result of the reform: immigrants who have stronger preferences for child quality over quantity are more likely to migrate to Germany after the reform, since their prospective children will enjoy better economic opportunities as German citizens. We will refer to this mechanism as the "composition effect". Second, couples that are close to the 8 years threshold might postpone their attempt to have a child in order to allow their future child to gain citizenship at birth. We refer to this as the "postponement effect". Finally, when making decisions about the total number of children, immigrants might take into account that future children will enjoy German citizenship if born once one of the spouses has fulfilled the 8 year residency requirement. While a priori it is not possible to make any conjecture on the size of these effects, the three of them should lead to a reduction in the fertility rates among the couples with less than 8 years of residence in Germany. Therefore, the difference between the fertility of immigrants with more than 8 years and those with less than 8 years would represent an effect of the reform that is biased towards zero.

In order to attenuate the "composition effect", we restrict the sample to households where at least one of the partners had arrived in Germany before 2000, whose decision to migrate to Germany was not affected by the citizenship reform. In total, there are 6,261 observations for the group of households where both partners have been for less than 8 years in Germany during the time period

between 1996 and 2005, as opposed to 26,070 observations for those households where at least one partner has been for more than 8 years in Germany. The two groups display large differences in the average number of years since migration to Germany of the partner who has been the longest in Germany: on average, 5 (with a s.d. of 2.19) years for those with less than 8 years, as opposed to 22 (with a s.d. of 7.7) years for those in the treatment group. Differences in the number of years since migration are associated with differences along many dimensions. For instance, women who migrated to Germany earlier in life are more likely to have attended school in Germany, a potentially important determinant of the fertility behavior.[1]

Since the number of years spent in Germany might lead to substantial differences in the fertility behavior, we restrict the analysis to the group of immigrants who are relatively close to the 8 years of residence threshold. We estimate the DD specification reported in equation 1 for two different samples: 1) immigrant households where the partner who has been the longest in Germany has been resident for a period between 1 and 16 years and 2) those where the partner with the longest spell has been in Germany for a period between 4 and 13 years. In Table AI we present immigrants' characteristics according to the time they have spent in Germany. Columns 1 and 2 report the characteristics of immigrants in households where the partner who has been the longest in Germany has been resident for less than 8 years and those in households where the partner who has been the longest in Germany has been resident between 8 and 16 years respectively. Columns 3 and 4 display the main characteristics of those in households where the partner who has been the longest in Germany has been resident between 4 and 7 years, and between 8 and 13 years respectively. As expected, smaller differences in terms of years spent in Germany are associated with smaller differences in terms of observable characteristics. In Table AII we present the estimates of equation 1 for the two subsamples. Columns 1 and 3 display the results for the specification with no controls, columns 2 and 4 report the results for our baseline specification. After the reform, immigrant households in the treatment group display a reduction in fertility compared to those households who have been in Germany for less than 8 years. When we use our baseline specification, the effect is 1.2 percentage points, irrespective of whether we focus on households where the partner who has been the longest in Germany has been resident for a period between 1 and 16 years or those where the partner with the longest spell has been in Germany for a period between 4 and 13 years. The coefficients are not statistically significant at conventional levels, but this is likely to be explained by the small number of observations. While the absolute size of the effect is perfectly in line with the one displayed in Table 2, its relative size is smaller since it represents 3.7% of the standard deviation of the fertility rate displayed in the sample.

[1]Results in Adsera and Ferrer (2011) show that women who migrated to Canada early in life display a fertility behavior that is not statistically different from Canadian women's, but the difference increases significantly with the age at migration.

TABLE AI—IMMIGRANTS' CHARACTERISTICS BY NUMBER OF YEARS SPENT IN GERMANY

	(1)	(2)	(3)	(4)
	8-16 years in Germany	0-7 years in Germany	8-13 years in Germany	4-7 years in Germany
	mean/sd	mean/sd	mean/sd	mean/sd
Age Mother	32.664	32.940	32.584	32.779
	(7.459)	(7.052)	(7.695)	(6.669)
Mother has primary education	0.196	0.174	0.195	0.186
	(0.397)	(0.379)	(0.396)	(0.389)
Mother has tertiary education	0.432	0.351	0.445	0.378
	(0.496)	(0.478)	(0.497)	(0.485)
Mother has tertiary education	0.336	0.437	0.323	0.397
	(0.473)	(0.496)	(0.468)	(0.490)
Mother selfemployed	0.016	0.014	0.017	0.016
	(0.126)	(0.116)	(0.128)	(0.124)
Mother unemployed	0.088	0.119	0.083	0.094
	(0.284)	(0.324)	(0.276)	(0.292)
Mother EU	0.217	0.192	0.222	0.168
	(0.413)	(0.394)	(0.416)	(0.374)
Log HH Income	2.570	2.320	2.534	2.206
	(1.251)	(1.675)	(1.226)	(1.368)
Total Number of Children	1.677	1.753	1.625	1.879
	(1.344)	(1.355)	(1.367)	(1.365)
Observations	622	730	483	511

Note: The sample includes households where the woman is in the age group 15-49. Summary statistics are computed using the 1999 survey. The dependent variable is a dummy equal to one if there is a child born within the last 12 months. In each column the number of years spent in Germany refers to the number of years spent by the partner in the household who has been longest in Germany.

Table AII—Results by the number of years spent in Germany

	(1)	(2)	(3)	(4)
	Newborn child			
	Between 1 and 16 years		Between 4 and 13 years	
Less than 8 years in Germany	-0.015*	2.677	-0.019	-2.536
	(0.008)	(3.545)	(0.011)	(5.008)
Less than 8 years in Germany*after	0.016	0.012	0.022*	0.012
	(0.009)	(0.009)	(0.012)	(0.011)
After	0.009		0.012*	
	(0.012)		(0.007)	
Time Dummies		x		x
Group Trends		x		x
Controls		x		x
Observations	13262		9408	
Mean of Dep. Variable	0.121		0.120	
SD of Dep. Variable	0.326		0.325	

Note: Robust standard errors clustered at the group-year level in parentheses. The sample includes households where the woman is in the age group 15-49. The treatment group is the reference category. The comparison group includes households where none of the parents has been longer than 8 years in Germany. The dependent variable is a dummy equal to one if there is a child born within the last 12 months. In Columns 1 and 2 the sample is restricted to households where the partner who has been the longest in Germany has been resident between 1 and 16 years. In Column 3 and 4 the sample is restricted to households where the partner who has been the longest in Germany has been resident between 4 and 13 years. Controls are the same as those included in Table 2 in the main text.

*** Significant at the 1% level.

** Significant at the 5% level.

* Significant at the 10% level.

Appendix B. Bounds on the obesity effect

We consider the relevant health outcome to be the outcome defined by the binomial variable that is equal to 1 if the child is classified as obese and zero otherwise. $o_{i,c}$ is then the average obesity rate of group i belonging to cohort c. To improve readability we use $c >= 2000$ to indicate all cohorts in our sample born in or after 2000 and $c <= 1999$ to denote all cohorts born before 2000. Given that we consider only children aged between 4 and 5 and that information on the weight and height of respondents is provided only by the 1999, 2003 and 2005 surveys, the only cohorts observed after 2000 are the cohorts of children born either in 2000 or in 2001. Let us now define two groups, T_1 and T_2, of sizes α and $(1-\alpha)$ respectively. The first is the group of children who would have been born anyway regardless of their citizenship rights; the second is the group of children who may not have been born were their prospective legal status to have been different. C is the comparison group. Note that in the following discussion we consider the two comparison groups defined in the previous part of the paper as a single comparison group.[2] The coefficient we observe is determined by the following expression:

$$\beta^{OBS} = [o_{T_1,c>=2000} - o_{C,c>=2000}] - [\alpha o_{T_1,c<=1999} + (1-\alpha)\, o_{T_2,c<=1999} - o_{C,c<=1999}]$$

while the true causal effect of the reform on obesity rates is

$$\begin{aligned} \beta^{CAUS} &= [\alpha o_{T_1,c>=2000} + (1-\alpha)\, o_{T_2,c>=2000} - o_{C,c>=2000}] \\ &- [\alpha o_{T_1,c<=1999} + (1-\alpha)\, o_{T_2,c<=1999} - o_{C,c<=1999}] \end{aligned}$$

It is straightforward to rewrite the two expressions as follows:

$$\beta^{CAUS} = \beta^{OBS} - (1-\alpha)\,(o_{T_1,c>=2000} - o_{T_2,c>=2000})$$

where $-(1-\alpha)\,(o_{T_1,c>=2000} - o_{T_2,c>=2000})$ is the correction we need to apply to β^{OBS} to obtain an estimate of β^{CAUS}. In the last expression $(1-\alpha)$ represents the percent decline in size experienced by cohorts 2000 and 2001 of treated children caused by the reform. In order to estimate $(1-\alpha)$ we need to estimate a specification that gives us the effect of the reform on these two cohorts. We then estimate the following specification, which is only a slight variation of our main specification:

$$\begin{aligned} Y_{it} &= \alpha + \beta_1 C_i + \beta_2 C_i * D_{t=1996\&t=1997} + \beta_3 C_i * D_{t=2000\&t=2001} + \beta_4 C_i * D_{t=2002\&t=2003} \\ &+ \beta_5 C_i * D_{t=2004\&t=2005} + \beta_6 C_i * t + \gamma X_{it}^{'} + \mu_t + u_{it} \end{aligned}$$

[2]Table AIII shows that results obtained from merging the two comparison groups are similar to the results obtained using our baseline specification.

The comparison group includes the two comparison groups described in Section II. C_i takes the value 1 if woman i belongs to a household where at most one parent is a foreign citizen. In addition, in place of a dummy for cohorts born after the reform, we use four different dummies, for survey years 1996-1997, 2000-2001, 2002-2003, and 2004-2005. $D_{t=2000\&t=2001}$ is a dummy variable that is equal to 1 if the survey year is 2000 or 2001.[3] β_3, the coefficient of the interaction term $C_i * D_{t=2000\&t=2001}$, is our coefficient of interest and can be interpreted as the effect of the reform on the fertility outcomes of women in the treatment group during the survey years 2000 and 2001. We can then estimate $(1-\alpha)$ by using the ratio of β_3 to the average of the dependent variable Y (the probability of having a child aged zero) among treated women before the introduction of the new nationality law, that is 0.08352. Table AIV in the Appendix provides an estimate of the coefficient β_3 which we can use to calculate $(1-\alpha)$ $((1-\alpha) = 0.0598)$. As already argued, we assume the highest possible bias, therefore $o_{T_2,c>=2000} = 1$. $o_{T_1,c>=2000}$ should instead correspond to the average obesity rates (0.1897) of the children in the treated group born after 2000. The size of the correction we need to apply then is $-0.0598(0.1897-1) = 0.048$ and the bounded estimates of our coefficient of interest, $\beta^{CAUS} = -0.08 + 0.048 = -0.032$.[4]

Table AIII—Merged comparison groups: Health outcomes

	(1)	(2)
	BMI	Obese
Comparison Group I & II*after	0.864*	0.080*
	(0.457)	(0.044)
Sample	Age 4-5	Age 4-5
Controls	x	x
Observations	10991	10991

Note: The treatment group is the reference category. The comparison group is the merge of Comparison Group I and II. Controls are the same as those in Table 6 in the main text.

*** Significant at the 1% level.

** Significant at the 5% level.

* Significant at the 10% level.

[3] $D_{t=1998\&t=1999}$ is the omitted time category employed in order to calculate the effect of the reform on fertility outcomes of the treated group in the survey years 2000 and 2001. This allows us to compare directly the 2000/2001 group with the group that was not affected by the reform but was born immediately before it was implemented. Also, this choice provides a higher estimate of the bias and therefore a more conservative estimate of the effect of the reform on the obesity rates among of immigrants' children.

[4] β^{OBS} is obtained by using the specification in equation 1 but merging the two comparison groups already introduced to define the comparison group. The results are provided in Table AIII.

TABLE AIV—TIME SPECIFIC EFFECTS

	(1)
	Newborn
Comparison Group I & II*96-97	0.002
	(0.003)
Comparison Group I & II*00-01	0.005
	(0.008)
Comparison Group I & II*02-03	0.006
	(0.004)
Comparison Group I & II*04-05	0.010**
	(0.004)
Time dummies	x
Controls	x
Observations	941586

Note: Robust standard errors clustered at the group-year level in parentheses. The sample includes households where the woman is in the age group 15-49. The treatment group is the reference category. The comparison group is the merge of Comparison Group I and II. The comparison group dummy is interacted with four different dummies, for survey years 1996-1997, 2000-2001, 2002-2003, and 2004-2005. The years 1998 and 1999 are the reference category. The dependent variable is a dummy equal to one if there is a child born within the last 12 months. Controls are the same as those included in Table 2 in the main text.

*** Significant at the 1% level.

** Significant at the 5% level.

* Significant at the 10% level.

Appendix C. Difference-in-Differences with propensity score method

In this section we present a further test to control for the possibility that our main results are the artifact of differential trends driven by the differences in the characteristics of the treatment and control groups. The analysis presented in this section uses as a control group only those households where there is one German parent. Households with one foreign parent are more likely to share characteristics - either observed or unobserved - with those in the treatment group. Second, in order to guarantee a higher comparability of the treatment and the control group in terms of observable characteristics, we combine the difference-in-differences method with a propensity score method (see Blundell et al. (2004) and Abadie (2005)). We first estimate two propensity scores, one for the assignment to treatment group in the pre-reform period and one for the assignment to the treatment group in the post-reform period. The propensity scores are estimated parametrically via logit models with a flexible specification that controls for mother's age, mother's education dummies, household income, state of residence dummies and dummies for whether the mother is German or a EU citizen. After restricting the sample to those observations for which the common support property holds, we estimate equation 1 using an Inverse Probability Weighting estimators, where

the weights are represented by the predicted propensity scores. In Table AV we show the results based on the main specification presented in column 5 of Table 2. For each outcome of interest we find results that are perfectly in line with those presented in the main section.

Table AV—Difference-in-Differences with Propensity Score

	(1)	(2)	(3)
	Newborn	At least one child 0-3	Number of children 0-3
Comparison Group I* after	0.012***	0.029***	0.032***
	(0.004)	(0.005)	(0.006)
Time dummies	x	x	x
Group Trends	x	x	x
Controls	x	x	x
Observations	76097	76097	76097

Note: Robust standard errors clustered at the group-year level in parentheses. Controls include mother's age, mother's education dummies, state of residence dummies and dummies for whether the mother is German or a EU citizen.

*** Significant at the 1% level.

** Significant at the 5% level.

* Significant at the 10% level.

Appendix D. Economic Shocks and Alternative Policies

In this section we investigate in deeper detail the possibility that we are capturing differential responses of households in the treatment groups to macroeconomic shocks and to policies that changed benefits and requirements associated with the welfare programs. In fact, the reduced immigrants' fertility might be the result of economic shocks or policies contemporaneous with the citizenship reform that had a differential effect on treatment group and comparison groups. The stability of the inflation and GDP growth rates in the period covered by our study does not provide support to the hypothesis that our results can be explained by macroeconomic shocks. Nevertheless, important changes in the labor market regulations and welfare programs' entitlements have been introduced in Germany after 2000.

Passed between January 2003 and January 2005, the Hartz Laws (I-V) were designed to tackle unemployment issues and attain a modern provision of services in the labor market. Among others, the measures implemented changes in the occupational training programs, subsistence payments, temporary employment, job search methods, and eligibility in unemployment assistance.[5] Figure 2 in the paper provides *prima facie* evidence that our results cannot be explained by the Hartz Laws since the drop in fertility among immigrants affected by the reform started in 2001, two years before the first set of measures was introduced.

In order to formally assess whether the reduced fertility can be explained by changes in female labor supply, we use equation 1 in the paper to test whether women in the treatment group changed their behavior in the labor market after the reform. Results based on our main specification are presented in Table AVI. We do not observe any change after 2000 in the probability that women in treated households (when compared to women in control households) are employed (column 1), and in the probability of working full time, irrespective of whether it is conditional on working or not (columns 2 and 3).[6] While employment outcomes might adjust slowly, women in the treatment group might intensify their job search and decide to postpone their fertility attempts. Columns 4 and 5 report the results of two specifications where the dependent variable is a dummy variable equal to one if the respondent declares to be looking for a job and the sample used consists of unemployed respondents (column 4) and already employed (column 5). Women in the treatment group seem to increase their efforts to look for a job compared to women belonging to mixed households, but not to those with only Germans. Overall, the results seem quite noisy and they are unlikely to explain the observed fertility drop. Changes in the generosity of the welfare programs might potentially impact fertility behavior of treated households. Nevertheless we do not find any evidence of differential changes in the probability

[5] See Farh and Sunde (2006) and Kluve and Jacobi (2006) for a detailed description of the Hartz laws and analysis of their effects.

[6] A woman is considered to be working full time if she reports to have worked 35 hours or more during the week before the survey.

that any household member is receiving a welfare benefit.[7] Taken together, these results do not support the hypothesis that the drop in fertility experienced by households in the treatment group after 2000 was driven by changes in the labor market outcomes of women.

The childcare system is a potentially important determinant of the fertility behavior. In Germany the eligibility criteria and the benefits of the childcare provisions do not differentiate between citizens and non-citizens. Nevertheless, their effect on female fertility might vary according to the socio-economic status of the household. A reform of the the childcare system was implemented in 2001, only one year after the nationality law reform. The 2001 law, however, was the last and probably the least decisive among a sequence of legislative acts aimed at increasing the length of parental leave and per child allowances.

The first policy reform in Germany was implemented in 1979 and increased paid maternity leave from 2 to 6 months. The second reform, in 1986, increased it from 6 to 10 months. Two legislative changes in July 1989 and July 1990 lengthened the job-protected leaves to 15 and 18 months, respectively. Finally, the reform implemented in 1992 increased it from 18 to 36 months, the longest in the world.[8] The 2001 law introduced only a minor extension of the set of choices available to parents by allowing both to take parental leave at the same time (with no change, however, in the total length of parental leave).[9] The 2001 law also extended the per child allowance to an additional small portion of the population: those households whose net annual income is between DEM 29,400 and DEM 32,200.[10] There is mixed evidence on the effect of child subsidies on fertility, but recent studies suggest that financial incentives stimulate fertility (e.g., Milligan (2005) for Canada; Cohen, Dehejia and Romanov (2013) for Israel). Therefore, the evidence in the previous literature seems not to justify a connection between the new childcare provisions and the sharp decline in the fertility of immigrants after 2001.

[7]The dependent variable is defined as a dummy variable equal to one if at least one of the parents within the household received at least one of the following benefits reported by the Microcensus survey: housing allowance, social benefits, unemployment benefits, support for education training/scholarships, nursing allowance, other public benefits.

[8]As noted by Schoenberg and Ludsteck (2007), there does not seem to be a long-run trend in German fertility rates between 1977 and 1993. Lalive and Zweimuller (2009), however, provide evidence that in Austria, extending paid parental leave from 1 to 2 years increased fertility.

[9]This is also unlikely to be relevant for the households in our treatment group as female labor force participation is lower among immigrants.

[10]Between US$ 21000-23000.

TABLE AVI—ALTERNATIVE EXPLANATIONS

	(1)	(2)	(3)	(4)	(5)	(6)
	Female is employed	Female works full time	Female works full time (conditional)	Female job search if unemployed if unemployed	Female job search on the job on the job	At least one HH member receives welfare benefit
Comparison Group I	2.343	-6.583*	-28.757***	-15.136***	-2.626	21.778
	(4.829)	(3.666)	(8.524)	(4.221)	(2.856)	(17.466)
Comparison Group II	-11.139**	-18.951***	-44.224***	-9.958***	3.731***	27.077
	(4.564)	(3.227)	(8.201)	(1.770)	(1.132)	(15.974)
Comparison Group I*After	-0.012	0.001	0.012	-0.032**	-0.008	-0.006
	(0.015)	(0.008)	(0.025)	(0.015)	(0.007)	(0.036)
Comparison Group II*After	-0.022	-0.011	-0.006	-0.003	0.005	0.002
	(0.014)	(0.007)	(0.024)	(0.006)	(0.003)	(0.032)
Time Dummies	x	x	x	x	x	x
Group Trends	x	x	x	x	x	x
Controls	x	x	x	x	x	x
Observations	941586	941586	672898	268688	672894	941586
Mean of Dep. variable	0.711	0.415	0.581	0.286	0.039	0.559
SD of Dep. variable	0.453	0.493	0.493	0.471	0.194	0.497

Note: Robust standard errors clustered at the group-year level in parentheses. The sample includes households where the woman is in the age group 15-49. Comparison Group I and II are defined as in Table 1. In column 1 the dependent variable is a dummy equal to one if the female member of the HH has been employed in the reference week. In columns 2 and 3 the dependent variable is a dummy equal to one if she has been employed full time. A woman is defined to be working full time if she has been working 35 or more hours in a usual week. In columns 3 the sample is restricted to women who declared to be employed. In column 4 the dependent variable is a dummy variable equal to one if the woman has been actively searching for a job while not employed. In column 5 the dependent variable is a dummy variable equal to one if the woman has been actively searching for a job while employed. In column 6 the dependent variable is a dummy equal to one if the partners within the household received at least one of the following benefits reported by the Microcensus survey: housing allowance, social benefits, unemployment benefits, support for education, training/scholarships, nursing allowance, other public benefits. Controls are the same as those defined in Table 2 in the main text.

*** Significant at the 1% level.

** Significant at the 5% level.

* Significant at the 10% level.

REFERENCES

Abadie, Alberto. 2005. "Semiparametric Difference-in-Differences Estimators." *Review of Economic Studies*, 72(1): 1–19.

Adsera, Alicia, and Ana Ferrer. 2011. "Age at Migration, Language and Fertility Patterns among Migrants to Canada." Institute for the Study of Labor (IZA) IZA Discussion Papers 5552.

Blundell, Richard, Monica Costa Dias, Costas Meghir, and John Van Reenen. 2004. "Evaluating the Employment Impact of a Mandatory Job Search Program." *Journal of the European Economic Association*, 2(4): 569–606.

Cohen, Alma, Rajeev Dehejia, and Dmitri Romanov. 2013. "Financial Incentives and Fertility." *Review of Economics and Statistics*, 95: 1–20.

Farh, Ren, and Uwe Sunde. 2006. "Did the Hartz Reforms Speed-Up Job Creation? A Macro-Evaluation Using Empirical Matching Functions." Institute for the Study of Labor (IZA) IZA Discussion Papers 2470.

Kluve, Jochen, and Lena Jacobi. 2006. "Before and After the Hartz Reforms: The Performance of Active Labour Market Policy in Germany." Rheinisch-Westflisches Institut fr Wirtschaftsforschung RWI Discussion Papers 0041.

Lalive, Rafael, and Josef Zweimuller. 2009. "How does Parental Leave Affect Fertility and Return to Work? Evidence from Two Natural Experiments." *The Quarterly Journal of Economics*, 124(3): 1363–1402.

Milligan, Kevin. 2005. "Subsidizing the Stork: New Evidence on Tax Incentives and Fertility." *The Review of Economics and Statistics*, 87(3): 539–555.

Schoenberg, Uta, and Johannes Ludsteck. 2007. "Maternity Leave Legislation, Female Labor Supply, and the Family Wage Gap." Institute for the Study of Labor (IZA) IZA Discussion Papers 2699.