CReAM
Centre for Research & Analysis of Migration

Discussion Paper Series

CDP 05/19

Lowering Welfare Benefits: Intended and Unintended Consequences for Migrants and their Families

Lars Højsgaard Andersen, Christian Dustmann, and Rasmus Landersø

Centre for Research and Analysis of Migration
Department of Economics, University College London
Drayton House, 30 Gordon Street, London WC1H 0AX

www.cream-migration.org

Lowering Welfare Benefits: Intended and Unintended Consequences for Migrants and their Families

Lars Højsgaard Andersen[a], Christian Dustmann[b], and Rasmus Landersø[a]

This Draft: March 2019

Abstract: Denmark's Start Aid welfare reform reduced benefits to refugee immigrants by around 50 percent for those granted residency after the reform date. The reform led to a sharp short run increase in labor earnings and employment, but it also induced a strong female labor force withdrawal, and a large and persistent drop in disposable income for most households. Furthermore, the reform caused a sharp increase in property crime among both females and males. Moreover, children's likelihood of being enrolled in childcare or preschool, their performance in language tests, and their years of education all decreased, while teenagers' crime rates increased.

Keywords: Social assistance, welfare state, labor market outcomes, migration.
JEL: E64, I30, J60

---

[a] ROCKWOOL Foundation Research Unit, Copenhagen

[b] University College London, Centre for Research and Analysis of Migration (CReAM) and ROCKWOOL Foundation Research Unit.

This paper benefitted from comments and suggestions by Joe Altonji, David Card, Raj Chetty, Steve Coat, Hillary Hoynes, Pat Kline and seminar participants at Berkeley, Chicago, Yale, Stanford, Cornell, ViVe, and the NBER Summer Institute. We are grateful to the ROCKWOOL foundation for funding this project. Christian Dustmann acknowledges funding from the DFG and the Norface Welfare State Future program.

1 Introduction

In response to the recent large migration flows and a sharp rise in anti-immigration sentiment, many governments are restricting access to welfare benefits for immigrants and their families. For instance, Canada took drastic measures to limit immigrants' access to social assistance in 2014 (following a first round of cuts in 2012), and Germany approved in 2016 a bill limiting access to social benefits for refugees not taking up jobs or training in low-wage sectors. The Netherlands proposed a large cut to benefits for newly arrived migrants in 2017, and recently Austria's coalition government proposed a policy making immigrants ineligible for the main minimum benefit during the first five years after entering the country while also capping benefit payments to refugees. Other countries currently discuss similar proposals for cutting benefits for immigrants and refugees.[1] The primary aim of these reforms is to incentivize employment and self-sufficiency, but in the absence of employment take-up the lower transfer levels reduce the household income of already low-income families, with potentially harmful effects on both adults and their children. As many of these reforms are recent (or in the implementation or proposal stage), and because data on affected immigrant populations are difficult to obtain, hardly any analysis exist that studies the immediate and longer-term consequences of such reforms on immigrants and their families.

In this paper, we overcome these challenges by analyzing the effects of Denmark's Start Aid welfare reform, which shares many of the features of more recent reforms or reform proposals. In 2002, the new right-leaning government introduced policies to reduce welfare benefits for refugees whose asylum claim had been approved after July 1st to around 50% of the previous "Social

[1] The policies implemented or proposed include restriction of immigrants' access to social assistance and public benefits in Canada (CBC, 2014), Finland, France, The Netherlands, Latvia, Lithuania (OECD International Migration Outlook 2017, 2018), Switzerland (Swissinfo, 2017), the Austrian governments' proposed cut in refugees' transfers, (Regierungsprogramm, 2017), and restrictions and further proposals in Germany (Library of Congress; Frankfurter Allgemeine, 2016). More generally, from 2000-2017, the EU27 countries passed 158 bills regarding refugees' and migrants' welfare eligibility, program requirements, or welfare levels (OECD International Migration Outlook 2006-2018, OECD Trends in International Migration 1998-2004).

Assistance" (SoA) level.[2] This new benefit is referred to as "Start Aid" benefits. Because of the very short reform implementation period (proposed in March, approved in June, and enacted on July 1 in 2002) relative to the lengthy asylum process, those granted residency around the cut-off date had already been in Denmark before reform proposal and approval. This creates a large discontinuity in transfer levels, which we use for identification.

Overall, while the reform led initially to an increase in employment and earnings, it also permanently decreased total income for refugees. In the first year after its implementation, average labor earnings doubled and employment rates increased from a pre reform rate of 10% to 19%, but total average income decreased by 40%. While the employment response fades over time, the reform also led to a sharp and more persistent increase in female labor force exits. This is explained by two features of the structure of welfare benefits. First, both pre reform SoA and post reform Start Aid are means tested at the household level. As the reform increases males' earnings and employment, the means test reduces females' transfer payments, and eliminates their incentive to stay in the labor force and attend integration programs, a pre-condition for benefit receipt.[3] Second, in couples where one spouse was granted residency before and the other was granted residency after the reform date, the first spouse was eligible for the generous SoA welfare benefits, while the second spouse was not eligible for any welfare benefits.[4] At household level, this led to the same transfers as for couples where both partners received residency after the reform (and both received Start Aid benefits). However, the asymmetric allocation of transfers within the household removes the incentive to participate in integration programs, and hence to stay in the labor force for the

[2] A refugee is an asylum seeker whose application for asylum has been approved. Upon receiving asylum, the individual is granted residency, and is entitled to welfare benefits.
[3] Integration programs involve on average 30 hours per week and consist of language courses, courses in Danish society and culture, and active labor market programs. Participation is a prerequisite for transfer receipt, and requirements commence immediately after residency.
[4] The majority of adult refugees are married couples (75% of our sample), and in most cases, the two spouses arrive in Denmark at different dates. Furthermore, couples' asylum cases are treated individually, and they receive residency at different dates.

partner who receives residency last (in 86% of cases the wife). Thus, both means testing and the asymmetric transfer allocation for couples who received residency on both sides of the reform date led to a sharp increase in females dropping out of the labor force by removing the penalty for non-participation in integration programs.

While the reform increased labor earnings for some households, it foremost induced a dramatic and persistent drop in disposable income for the residual 70% of households that still relied entirely on welfare benefits. Studying the consequences for individuals and their families, we show that the reform led to a sharp increase in property crime (such as grocery shoplifting) among adult males and females. Furthermore, the reform also produced adverse effects on children. In particular, children who receive residency at preschool age become less likely to be enrolled in childcare or preschool, and more likely to be among the poorest performers on language tests. Moreover, those who receive residency after the reform in their teens complete eight months less education, an effect driven mainly by boys, who are more likely to leave education and enter unskilled youth employment or welfare reception. Finally, the reform sharply increases youth crime for both boys and girls.

Our analysis provides critically needed evidence on the immediate and longer-term effects of welfare reforms targeted at immigrants or refugees, similar to those that are currently being implemented or discussed in many Western countries. Moreover, by studying the labor supply responses to a reform that drastically reduced means tested transfers for refugees, we add to the literatures that analyze labor supply responses to means tested transfers (Eissa and Liebman, 1996; Eissa and Hoynes, 2004; Hoynes, 1996; Meyer and Rosenbaum, 2001; Saez, 2002; Moffitt 2002, 2015, Mogstad and Pronzato, 2012), the nexus between extensive and intensive margin responses (Heckman, 1993; Blundell and MaCurdy, 1999), and welfare transfers to immigrants (e.g., Borjas, 2002).[5] Furthermore, the reform we analyze allows us to investigate incentive effects of two distinct reform implementation regimes in which couples encounter equal transfer reductions at the household level, but different reductions at the individual level. Our finding that labor supply elasticities differ markedly between the two design regimes adds to the large literature identifying labor supply elasticities (see e.g., Chetty et al., 2011a; Kleven and Schultz, 2014; Saez et al., 2012), as our results suggest that estimated labor supply elasticities may be highly sensitive to how policies are implemented, and the specific populations they target.

We also speak to the broad literature on family income, crime, and child outcomes. While previous studies have associated crime with the timing of welfare payments (e.g., Foley, 2011; Carr and Packham, 2017), welfare eligibility of criminal offenders and recidivism (Yang, 2017), and state variation in welfare reform implementation and crime levels (Corman et al., 2014), we provide direct causal evidence of benefit levels' effects on crime for a disadvantaged population group. We show that the reform, and the resulting income reduction, induces adults and adolescents of both genders to commit significantly more crime.

In addition, we show that the transfer reductions also have serious consequences for the target group's children in dimensions such as language acquisition and education. Our findings thus add to the literature linking family income, welfare reforms, and tax credits to children's outcomes (e.g., Aizer et al., 2016; Chetty et al., 2011b, Dahl and Lochner, 2012; Hoynes et al., 2016; Løken et al., 2012; Løken et al. 2018).

The remaining sections unfold as follows: Section 2 describes the Start Aid reform as well as the data and the sample. Section 3 details our identification strategy. Section 4 presents the

[5] Bratsberg et al. (2010) study trends in immigrants' welfare reception and employment after residency to Norway. In the context of Denmark, Huynh et al. (2007) and Rosholm and Vejlin (2010) are early studies of the Start Aid reform's immediate average impact on employment and unemployment length in a duration analysis, respectively. See also Moffitt (2002; 2015), and Chan and Moffitt (2018) for reviews of the literature on means tested transfers, and Krueger and Meyer (2002) for an extensive review of the literature on social transfers as UI benefits.

primary effects of the reform on labor supply and income. Section 5 presents the reform's secondary effects on crime and children's educational outcomes. Section 6 concludes the paper.

2 Background and Data

2.1 Social Assistance, the Start Aid Reform, and Benefit Eligibility

Denmark's Social Assistance (SoA) benefits are among the world's most generous and the country used to have one of the most liberal refugee immigration laws (Andersen et al., 2012; Huynh et al., 2007; Pedersen, 2010). Following large inflows of individuals with high levels of welfare uptake, net welfare transfers to non-Western immigrants reached 0.83% of GDP and 3.4% of total public spending in 2001 (Matthiessen, 2009). On March 1, 2002, a recently elected Danish government proposed a bill that replaced SoA for refugees with a new benefit scheme – "Start Aid" – in order to promote their labor market participation (Danish Prime Minister's Office, 2002).[6] Approved on June 6, and implemented on July 1, the reform assigned refugees granted residency after the reform date to Start Aid, where transfers are approximately 50% lower than SoA transfers (rates are based on the individual's age and family type, see Table A.1).[7]

To receive residency, refugees first have to request asylum in Denmark. Most do so after having entered the country as undocumented migrants. Once asylum is requested, the applicant is transferred to a specific reception center (Sandholmlejren, a set of former military barracks). Once the formal application process begins, the Danish Red Cross assigns the applicant to an

[6] "To ensure that refugees and immigrants living in Denmark are better integrated and find employment more quickly, the incentives for finding employment must be strengthened" (author translation of official remarks: http://webarkiv.ft.dk/Samling/20012/lovforslag_som_fremsat/L126.htm, accessed 03-20-2017). Appendix B.1 provides a more detailed description of the background and the Start Aid reform

[7] Going from SoA to Start Aid, the reform brought benefit levels for refugees down on par with median "Temporary Assistance for Needy Families" and "Supplemental Nutrition Assistance Program" levels in the U.S. (TANF: Falk 2014 who reports levels by state and year; SNAP: http://kff.org/other/state-indicator/avg-monthly-food-stamp-benefits).

accommodation center (refugee camp) while the application is processed by case workers at the Danish Immigration Service. Food (either served or via food stamps) and health care is supplied in the centers, but individuals are not allowed to work before the final approval of residency. For identification, we exploit the fact that the application process takes 15 months on average (Hvidtfeldt et al., 2017), which effectively randomizes individuals (who are already in Denmark) to Start Aid (vs. SoA) treatment according to when they are granted residency around the reform's implementation cut-off date. There is no cap on the number of residencies Denmark grants within a specific period, which could otherwise affect the number and timing of approvals. We describe our identification strategy in more detail in Section 3.[8] Upon being granted residency, refugees are assigned to a municipality. Importantly, both SoA and Start Aid eligibility is conditional on participation in an "integration program", which comprises language courses, courses on Danish society and acculturation, and active labor market program participation.[9] Refugees who fail to comply with these obligations immediately become ineligible for benefits (SoA and Start Aid).

We categorize individuals who are not working but participating in integration programs and are available to the labor market as "unemployed". Individuals who are not in work and do not participate in integration programs are ineligible for transfers, and we categorize these as "dropped out" because they have effectively left the labor market. A third and minor group of individuals are those who receive disability benefits and are exempt both from integration programs and the

[8] Married applicants are each assigned their own asylum case ID and processed individually even if they apply together on the same date. In our sample, 26% of the married couples have the same application and approval dates, 9% have different application dates but the same approval date, and 65% share neither the application nor the approval date. Unmarried couples are processed as two single individuals having independent case processing times.
[9] The integration program consists of two main components. The first is mandatory Danish and cultural lessons several times per week, and the second is active labor market programs. Within the first week after residency, each refugee receives a schedule by the municipality, with specific guidelines for course participation, etc. This schedule is revised every quarter. The two components together take up about 30 hours per week. The obligation to comply with the integration program discontinues during employment spells but resumes in case of new unemployment spells. See Law of Integration of Immigrants in Denmark: https://www.retsinformation.dk/Forms/R0710.aspx?id=28907#K4

transfer reduction. The "dropped out" individuals and those who are on disability benefits are together categorized as "not in labor force (NILF)".

Couples, the focus of our analysis, can be classified in three groups (Figure 1). If both spouses receive residency before the reform, both are entitled to SoA (Panel A). If both receive residency after the reform, both are entitled to Start Aid (Panel B, "Type A" couples). If one spouse receives residency before, and one spouse after the reform (Panel C, "Type B" couples), the first spouse keeps the full SoA, while the last arriving spouse's (LAS) benefits are capped such that total household benefits do not exceed that of two spouses on Start Aid (i.e. the same household level transfers as Type A couples receive). Because SoA is approximately twice as high as Start Aid, the LAS is effectively not entitled to any benefits. That also means that the spouse who receives residency last, receiving no benefits, cannot be penalized for non-participation in integration programs, a circumstance that is important for our findings below.

SoA and Start Aid are both means tested at the household level (with eligibility evaluated monthly) such that any labor earnings from the first earned dollar and onwards reduce benefits of both partners. Hence, refugees lose not only their own SoA or Start Aid due to their labor earnings, but their spouses' benefits are also cut. Means testing thus works as a household level Negative Income Tax that provides strong extensive margin disincentives (Immervoll et al., 2007; Saez, 2002). Importantly, those who lose benefits due to means testing are still recorded as "unemployed" as long as they meet the integration program requirements. Once they stop meeting these requirements, their status changes to "dropped out". Being the lowest tier of the Danish welfare system, there are no time limits on SoA and Start Aid receipt as long as recipients satisfy the eligibility criteria (participation in integration programs).

Figure A.1 describes the means testing. It shows the conversion of labor earnings into household gross incomes (on a monthly level) and the implied marginal tax rates from means testing for couples, where we distinguish between pre reform, Type A, and Type B couples (solid, dotted and dashed lines, respectively).[10] The vertical difference between the solid and the dotted / dashed lines in the intersections with the y-axis in Panel A of Figure A.1 shows the monthly benefit reduction induced by the reform. The slopes represent the means testing rates, and the resulting implied marginal tax rates are thus 1 minus these slopes (see Panel B of Figure A.1).[11] The means test implies a marginal tax rate of between 83% and 100% from any earnings above zero up to a break-even point at which SoA or Start Aid are fully exhausted and $1 earned increases gross income by $1. For pre reform couples the break-even point is at around $3,000 per month, while it is at around $2,000 and $1,500 for Type A and B couples respectively.[12] Thus, both pre and post reform the implied marginal tax rates lie close to 100% for the ranges below the break-even points, and at the usual low bracket income tax of 44% for labor earnings above the break-even point (Panel B of Figure A.1).

2.2 Information about incentives

In order to respond to the benefit reduction and changed incentives, refugees need to be aware of means testing and household type-dependent variations. Moreover, they need to understand how noncompliance with integration programs affects benefit eligibility and when benefits are capped. These issues are explained to potential welfare beneficiaries by the municipality of residence in

[10] Because 91% of couples have children, we use a one-child family as benchmark for couples' transfers. Table A.1 shows the extensive margin implied marginal tax rates and the break-even points by family type. Transfers are subject to the same income taxes as labor earnings, which is factored in when calculating marginal tax rates in Fig A.1B.
[11] Figure A.1 presents labor earnings and break-even points based on employment at the average collective bargain minimum wage. While Denmark has no formal minimum wage, the highly unionized labor market has collective bargain minimum wages, which serve almost universally as minimum wages.
[12] The break-even point for Type B couples is lower as it combines pre- and post-reform features. While total household income when the spouses are not working is the same as if both were on Start Aid (Type A), all household transfers are paid as SoA to one spouse. Therefore, the discount to means testing equals the pre reform SoA discount, resulting in a break-even point on a monthly level that is around $500 lower than for Type A couples.

physical meetings (with an interpreter when required) and in writing (see Appendix B.2 for samples of such letters of notice).[13] This implies that refugees are made aware of the particular incentive structure they face, including i) compulsory participation in an integration program, ii) the withholding of transfers for noncompliance, and for Type B couples, iii) that only the spouse who receives residency first will receive transfers, which makes ii) not applicable from the perspective of the LAS.

2.3 Data and Samples

Our analysis is based on register data recorded by public agencies and then compiled and organized by Statistics Denmark. The data contain unique personal identification numbers for individuals, their spouses, and their parents, which allow us to merge individuals to their entire family. Our sample consists of refugees that have been granted residency.[14] The exact date of residency defines each refugee's treatment status (pre vs. post reform) as well as our running variable (the time between residency date and the reform). Using the individual identifiers, we add information on labor market outcomes, occupation, income, education, crime, age, gender, and date of birth.

We study individuals aged 18 to 55 when granted residency, collect additional data on their children, and restrict our observation window to those granted residency between January 1, 2001, and December 31, 2003. From this initial sample of 8,512 individuals, we exclude refugees from Afghanistan (1,833) and the former Yugoslavia (932) because contemporaneous conflict-induced changes in inflow from these two countries could compromise the balancing of the sample around the reform. We also exclude those who re-migrate within the following nine years (in Section 4.1

[13] Danish authorities are required by law (Administrative Law, section 7, no. 1) to ensure that citizens have understood the rules and regulations they face when receiving benefits and when there are changes to their entitlements.
[14] Our sample only includes refugees and individuals who are family reunified with refugees, and not labor migrants, their families, and other non-refugees, who are not eligible for SoA or Start Aid and thus unaffected by the reform.

we test for selectivity over the reform period to ensure that re-migration patterns are not related to the reform).[15] Our *base sample* thus consists of 4,843 individuals who receive residency within our observation window, aged 18 to 55 at their residency date. Collectively they had 3,299 children aged 0 to 17 at the time residency was granted. As the focus of our analysis is on couples, we add the spouses of all individuals in the base sample. This results in a balanced *couples sample* of 4,072 individuals from 2,036 couples (57% of couples both receive residency pre reform, 13% both receive residency post reform, and 30% receive residency on both sides of the reform).[16]

2.4 Outcomes

In our analysis of reform effects on income, we consider four measures, all based on tax authority records: labor earnings, used to identify employment-related responses to the reform; transfer income, used to investigate the reform's direct effect through welfare benefit reduction; and total gross income and net-of-tax (disposable) income, used to measure the reform's total effect on income. We supplement this with data on hourly wage rates.

Each individual's labor market status is measured from the first full year after residency onwards, and we distinguish between three mutually exclusive states: *employment*, *unemployment*, and *not in the labor force* (NILF). As we discuss above, NILF comprises those who are permanently unable to work (on disability benefits), and those who are non-employed but fail to meet transfer eligibility requirements by not attending the integration program and thus have

[15] As noted earlier, the lengthy asylum process (on average 15 months for our sample) excludes the possibility that announcement effects compromise our identification, because those receiving residency around the reform's enactment date had already submitted their application before the reform was proposed. Similarly, a contemporaneous bill that changed the rules governing when (but not if) individuals could apply for permanent asylum (see Kilström, Larsen, and Olme, 2018) does not affect our identification, as it took effect only for asylum applications that were lodged from March 2002 onwards.

[16] All results we report below are robust to limiting the sample to the 90% of the couples sample where both spouses are granted residency in the ± 18 month window. Our sample of children does not include unaccompanied children, and the children in our sample are always assigned residency at the same date as one of their parents.

"dropped out" of the welfare system. We further categorize occupations (if employed) based on type and skill intensity.

In our analysis of "unintended" reform effects we study crime for adults as well as those who receive residency as teenagers with their parents. The crime data are based on police and court records for all crimes in Denmark. We observe offense dates, charges, arrests, incarcerations, and convictions for our entire sample. In our analysis of the impact the reform has on crime, we report results for convictions, which are categorized by detailed codes, allowing us to identify the exact crime type committed.

We also study children's education outcomes using register data on childcare or preschool attendance, Danish test scores in primary school (and the three subsets underlying the aggregate score: language comprehension, decoding, and reading comprehension), and total years of completed schooling.[17] We further investigate teenagers' education enrollment and labor market attachment. All outcomes and data sources are described in detail in Appendix B.3.

2.5 Descriptive Statistics

Figure 2 plots the distribution of average gross income from 2003-2007 for adult refugees receiving residency in 2002 by whether they received residency before or after the reform date, along with the income distribution of the full population. While refugees receiving residency before the reform date are clustered in the lowest 15 percentiles of the Danish income distribution with average annual incomes of around $15,000 or below, almost all refugees receiving residency after July 1st 2002 are placed in the lowest 5% of the income distribution with average incomes

[17] *Language comprehension* is the ability to construct the meaning of spoken language. *Decoding* is the ability to relate text (letters and words) to the sounds and meaning of the spoken language. *Reading comprehension* is the broader construct that is based on the two former subsets.

below $10,000. Thus, while among the poorest groups in Denmark before the reform, refugees become the group with the lowest income after the reform.

Panel A of Table 1 lists the covariate means for the base sample of adults aged 18 to 55, distinguishing between those who received residency before and after the reform. Fewer residencies are granted post reform than pre reform as the inflow of refugees to Denmark shows a downward trend over our sample period. Of the refugees in the base sample, 80% are immigrants from predominantly Muslim countries and around half have Iraqi origin. Residency based on refugee status is granted to 62% of the sample, whilst the remainder receive residency as a result of family reunification. Each adult has on average two children. Although the table reveals some differences between the pre and post reform groups (e.g., in the share of females), it should be noted that the key assumption in our setup is comparability in the limit around the reform's cutoff date. Our tests of this assumption (Section 4.1) show that observable characteristics are balanced, and that there are no discontinuities in covariates around the timing of the reform.

Panel B of Table 1 reports the covariate means for all couples where at least one spouse receives residency within our sample window, by gender and whether residency was granted before or after the reform. Demographic characteristics of the pre and post reform groups are very similar, although the female share is higher in the latter. This gender difference is a result of the downward trend in refugee inflow coupled with male refugees typically arriving first, followed by wives and children. Males are the first to receive residency in 86% of couples and therefore receive residency by refugee status rather than on family reunification basis, which is most common for females. The means of the remaining characteristics are similar across the groups except for country of origin, where differences stem again from trends in refugee inflow.

3 Estimation and Identification Strategy

The reform we study induced a large drop in transfers for the refugees who were granted residency after the reform's cut-off date. Based on this, we estimate the effect of the reform using a regression discontinuity design that compares refugees who receive residency just before and just after the reform date:

$$y_{i\tau} = \alpha + \beta * reform_i + g(Z_i)'\pi + X'_i\gamma + \epsilon_{i\tau}, \quad (1)$$

where $y_{i\tau}$ is an outcome for individual i measured τ years after residency is granted, $reform_i$ is a dummy variable indicating whether individual i receives residency after the reform date, $g(Z_i)$ is a running variable counting months between the date of residency and the reform date and allowing for separate trends on each side of the reform. The vector X_i measures observable characteristics, and $\epsilon_{i\tau}$ is an idiosyncratic error term.[18] We also analyze heterogeneous effects by interacting $reform_i$ and $g(Z_i)$ with elements of X_i.

The key parameter of interest is β. It measures the effect of being eligible for Start Aid instead of SoA among those individuals granted residency just around the cut-off date. It takes on average 15 months for individuals in our sample before the residency decision is made. Hence, those who receive residency close to the reform implementation date have arrived in Denmark before the reform's announcement, which was just 4 months earlier.

Focusing on couples, the treatment state depends on both spouses' timing of residency, and we therefore have two post reform treatment categories, as detailed in Section 2. To capture this, we extend Equation (1) to allow the outcome of individual i in household f to be affected by one's own timing of residency and that of one's spouse. We define three states. First, our baseline state, where both spouses receive residency before the reform and hence both are eligible for full SoA.

[18] We cluster standard errors by the running variable. In practice, we define g() as two linear functions before and after the reform. In Table A.2, we replicate our main results using alternative specifications.

Second, both spouses receive residency after the reform and are thus both eligible for Start Aid (Type A). Third, spouses receive residency on each side of the reform and have their benefits capped at two times Start Aid (which roughly equals SoA), paid to the spouse who receives residency before the reform, such that the spouse who receives residency after the reform is effectively made ineligible for any benefits (Type B). We define Type A and Type B couples by two non-overlapping treatment dummies, A_i and B_i, with baseline couples as the reference category.[19] We estimate the reform's effects on outcome $y_{i\tau f}$ of individual *i* from family *f* as:

$$y_{i\tau f} = \alpha + \beta_1 * A_{if} + g(Z_{1f})'\pi_1 + \beta_2 * B_{if} + g(Z_{2f})'\pi_2 + \epsilon_{i\tau f} \qquad (2)$$

where $g(Z_{1f})$ and $g(Z_{2f})$ control for the running variables that count the months to and from the reform for each spouse while allowing for different trends pre and post reform. The parameters β_1 and β_2 measure the effects for Type A and B couples, respectively, where baseline couples are the reference category. We also interact Equation (2) with gender, thereby estimating β_1 and β_2 separately for males and females.

4 The Primary Effects of the Start Aid Reform

Section 4.1 provides balancing tests that establish the validity of our design. Section 4.2 presents the reform's average effects on income and labor market outcomes for all refugees for the first five years after residency. Section 4.3, focusing on couples (75% of the adult sample), investigates the incentive effects of the means test and the two different reform implementations (for Type A and Type B couples). Section 4.4 reports the reform's long run labor supply effects.

[19] See Card et al. (2007a) for a further discussion of identification with double discontinuity.

4.1 Balancing Tests

Our key identifying assumption is that it is as good as random whether residency is granted right before or right after the reform's cutoff date. Table 2 reports estimates from regressing a dummy indicating whether the individual receives residency before or after the reform on the covariates from Table 1 and the running variable, and tests for joint significance of the covariates in these regressions to assess whether observable characteristics change around the reform date. We conduct this test for the base sample of adults (Panel A), the full sample including children (Panel B), and the couples' sample (Panel C). There are no jointly significant differences across the reform date, with p-values between 0.359 and 0.794.

We report additional balancing tests in the Appendix. For a visual balancing test around the reform date, Figure A.2 shows for each value of the running variable employment, unemployment, and NILF rates during the first year post residency predicted from an OLS regression using the same covariates (following Card et al., 2007a). The pre and post reform slopes are connected with no discontinuities in the predicted outcomes at the reform date, indicating no compositional changes to the sample around the cutoff.

Table A.3 shows estimates from regressions of each covariate separately on the reform dummy. Column A adds a variable measuring the exact time individuals spent in refugee camps before being granted residency (Hvidtfeldt et al., 2017), to test directly whether waiting time before residency changes following the reform. To test for selectiveness of remigration, we expand our sample in Column B by adding those who re-migrate and include a dummy that measures whether the individual leaves Denmark over the next 9 years. There is no evidence for selective remigration. Column C adds children to the base sample, and Columns D uses the couples sample and tests for selectivity in who is the first and last arriving spouse. Overall, in only one of the 37 individual balancing tests is the estimated parameter significant.

One additional concern may be that case workers respond to the reform by granting more residencies just before the reform date. Figure A.3 presents McCrary tests of differences in the running variable density (residencies per month) around Start Aid implementation using both optimal bandwidth selection, and smaller and larger bandwidths to confirm robustness. None of the specifications reveal any structural breaks.

Finally, we run placebo tests where we estimate Equation (1) using placebo reform dummies that vary from five months before the reform to five months after (we also change the associated running variables and the ± 18 months sampling window). Panel A of Figure A.4 reports the t-values from the estimated β's using transfer income (for the full sample and by gender), while Panel B uses employment, unemployment, and not in the labor force as outcomes. The graphs show that the t-values are between 0 and 1 for placebo reforms more than four months on either side of the actual reform, but increase as the timing of the placebo reforms converge towards the actual reform date and reach their maximum level when the "true" reform date is used (center of the figure). Thus, the estimated effects of the reform we report below are not a result of random fluctuations in the outcome variables.

4.2 Individual Level Effects of the Reform

As a first illustration of the reform's impact, Figure 3 shows that log total gross income averaged over five years drops around the reform date by approximately 0.4 log points (or 45%), while net of tax income declines by slightly less (40%), suggesting that the reduction is only modestly cushioned by tax progression. Hence, for the average individual in our sample, Start Aid resulted in a large and persistent fall in disposable income.[20]

---

[20] Figure A.5 shows that pre reform in the first year after residency only 10% of income comes from labor earnings (about $1,900) and the remaining 90% comes from transfers (about $20,500). The figure also shows that, while labor

In Panels A-B of Table 3, we separate the effects on transfers and labor earnings by time since residency, using levels instead of logs due to the presence of zeros in annual individual income, and report the pre reform means as benchmarks in the first column of each panel. Annual transfer income drops by approximately $10,000, $8,000, and $5,000 in years 1, 2, and 3–5 after residency, corresponding to 55%, 50%, and 30% reductions, respectively. At the same time, labor earnings rise by $1,100 – $1,560, amounting to increases of 60%, 35%, and 15% in years 1, 2, and 3–5, respectively. Yet, these increases are far from compensating for the lower benefit levels.

Panels C-E of Table 3 present the effects on employment, unemployment, and NILF for the first five years after residency. The first pair of columns show that average first-year employment rates post reform almost double, from 10.3 percent to 20 percent. In year 2, employment effects remain sizeable relative to the pre reform mean (at 7 percentage points, or 37%), but drop to about 13% in years 3-5. Interestingly, the reform lowers unemployment by around 17, 16, and 10 percentage points in years 1, 2, and 3–5, far more than the increase in employment, a difference that is explained by increases in the NILF rates.[21]

To investigate the increase in the fraction not in the labor force, we next provide in Table 4 estimates by gender, and distinguish between "not in the labor force" and the subset "dropped out", due to withdrawal from integration programs.[22] The estimates show that employment effects in the first two years are driven by males whose employment uptake is accompanied by a corresponding drop in unemployment. Females, in contrast, show only a small and insignificant employment response initially, but experience a large reduction in unemployment rates, which

earnings increase as a response to the reform, total gross income drops to almost half of the pre reform level and average net-of-tax income by around 40% to (or below) the estimated subsistence minimum in Denmark.

[21] The sharp discontinuity around the reform date is further illustrated by Figure A.6, which shows labor market outcomes during the first full year since residency around the reform date. Employment rates increase from a pre reform mean of 10% to around 20%, while unemployment rates decrease from 90% to around 70-75%, with the difference due to an increase in the NILF rate.

[22] The residual between all "not in the labor force" (NILF) and the "dropped out" are individuals on disability benefits, as mentioned earlier, and constitute 0.7% of the sample. These individuals are not affected by the reform, and responses of that group to the reform are virtually zero and insignificant.

coincides with a large increase of permanent withdrawals from the Danish labor (Panel D). These withdrawals are not driven by the low educated but are of similar magnitude for those with less and more than 12 years of education, see Panel D in Table A.4. Thus, while increasing employment for males, the reform led to a large rise of females exiting the labor force. These exits are far more persistent than males' employment response: while male employment effects decrease to about 4 percentage points (and statistically insignificant), female dropout rates are still 10 percentage points higher. Thus, the large drop in unemployment is only partly explained by employment take-up, a finding that underscores the importance of distinguishing between welfare benefits' effects on unemployment versus total non-employment and employment, as stressed by Bratberg and Vaage (2000), Card et al. (2007b) and Kyyra and Ollikainen (2008).

4.3 Household Level Responses and Reform Designs

We next investigate the mechanisms behind the reform's effects on employment and females' labor force exits, focusing on couples who constitute 75 percent of individuals in our adult sample.

4.3.1 Employment, Disincentive Effects, and Nonparticipation

Figure 4 presents the intuition behind the reform's simultaneous effects of increased labor supply and labor force drop outs for responses of Type A and B couples in a simple static labor supply framework, where the horizontal axis carries leisure/work and the vertical axis carries income.

Consider first the labor supply decision for Type A couples. In Figure 4A the dashed lines represent the pre-policy and post-policy budget sets (see also Figure A.1A). The almost horizontal parts of the budget sets correspond to the range of labor supply affected by means testing, while the slope of the steeper segment (after the kinks at the break-even points, where labor earnings equal SoA/SA payments) corresponds to the wage w.[23] The way we have drawn the figure, it is optimal for the pre reform couple to draw transfers and attend integration programs rather than supplying labor (point X). The reform reduces non-labor income (X to Y1), by reducing SoA for each partner to Start Aid. It also decreases the implicit return to sitting integration courses, determined by the ratio of transfers to required hours of integration program attendance (indicated by the change in the slopes of the grey dotted lines from X-F to Y1-F). As a response to the reform, the couple will now move from Y1 to Y2, supplying some labor at wage rate w and thereby reaching a higher indifference curve.

Consider now the same initial situation for Type B couples (Figure 4B). Again, transfers drop by the same amount, moving couples to Y1. Contrary to Type A couples, this transfer reduction is achieved by maintaining SoA payments for the first arriving spouse and eliminating payments to the LAS. Consequently, the LAS cannot be penalized for dropping out of integration programs, so Type B couples can increase household level leisure without reducing transfer income, by moving from Y1 to Z. As the figure suggests, the price of leisure over the range Y1-Z is zero, and the implicit return to sitting integration courses couples can realize by moving to point Z goes back to the same rate as for pre-reform couples (slope of the line Z-F).

Figure 4C shows how means testing influences the reform's effects on employment and labor force participation. We have drawn the post reform situations for Type A and Type B couples as points Y2 and Z. Suppose now there is no means test, a situation indicated by the grey solid line with slope w through point Y1. In that case, both Type A and Type B couples can move to a higher indifference curve by supplying labor at point Y3. This illustrates that the means test dampens the

[23] We treat the household as unitary maximizing husbands' and wives total utility as a function of the sum of their leisure and consumption (from income or transfers), assuming convex preferences and an upward sloping labor supply curve. See Lemieux and Milligan (2008) for a similar illustration. While the figure describes the implicit taxes from means testing, we abstract from income taxes.

employment response to the transfer reduction and removes incentives for females to participate in the labor force.

Figure 4 predicts that the reform increases employment, that this effect is largest for Type A couples, and that labor force exits of females should be most prevalent in Type B couples. The estimates in Table 5 confirm these predictions.[24] The table presents estimates where effects on labor market outcomes are separated by gender and household type (Panel A and B).[25] The table shows that male employment in year 1 increases by 15 and 8 percentage points in Type A and Type B couples respectively, and by roughly 10 percentage points for both household types in year 2. Female employment in Type A couples increases by 8 percentage points, unemployment rates decrease by 17 and 11 percentage points in years 1 and 2, and dropout rates increase by 8 and 6 percentage points. In contrast, females in Type B couples have a more muted (and insignificant) employment response of about 3 percentage points, but a far larger reduction in unemployment of more than 20 percentage points, which is accounted for by an increase in the drop-out rate. Thus, as suggested by Figure 4, the reform induces a larger initial labor supply response in Type A couples than in Type B couples and makes females in Type B couples 8 (15) percentage points more likely to exit the labor force in year 1 (year 2) post residency than in Type A couples.

Not attending integration courses and lacking basic language skills likely harm both females' longer run integration, and the language development and prospects of their children (Hoff, Laursen, and Tardiff, 2012). However, it is not clear which female dropouts can be attributed to the reform design – and specifically the reduction in benefit levels – and which arise from the household level means test, which means both partners have their benefit entitlement cut when

[24] The reform's effects on singles correspond qualitatively to the average effects for males (Table 4) with employment effects around of 13-15 percentage points in year 1 and 2 (results available upon request).

[25] Due to different residency dates, "time since residency" may now capture different periods for each spouse. We therefore align spouses' outcomes by defining "time since residency" as time since the LAS's residency. To ensure that results are not driven by the "time since residency" definition, Appendix section B.4 replicates our findings for labor market outcomes using time since first arrived spouse's residency.

one partner's earnings increase. We cannot address this in a causal way, as our reform only identifies the overall reform effect. To get some sense of the relative magnitude between couples' responses, we can proxy the "direct" effect of the reform on female drop outs, net of the effect working through changes in husbands' labor supply, by controlling for spouses' labor supply and how it changes across the reform when we estimate Equation (2).[26]

Column 5 in Table 5 shows that the estimates of the reform's impact on "dropped out" for females in Type A couples, once we "eliminate" the impact of means testing, are insignificant and close to zero. However, for females in Type B couples, the probability of dropping out is still about two-thirds the magnitude of the overall dropout rate, reducing from 15 to 10 (20 to 14) percentage points in year 1 (2). Hence, means testing is likely to explain all labor force exits in Type A couples, but less than half in Type B couples, with the remaining exits likely arising from the design feature of the transfer allocation scheme for Type B couples.

Panel C of Table 5 reports the implied elasticities of employment with respect to benefits for Type A and B couples, measuring percentage change in employment from a one percent change in benefits.[27] Elasticities differ substantially between the two types, with females' employment elasticities in Type A couples being more than three times larger than those in Type B couples

[26]Using "dropped out" as outcome $y_{i\tau f}$, we define a new dummy variable $Emp_{i\tau f}$ which is equal to 1 if one or both spouses are in employment, and we estimate the following expanded version of Equation (2):

$$y_{i\tau f} = \alpha + \ddot{\beta}_1 * A_{if} + g(Z_{1f})'\pi_1 + \ddot{\beta}_2 * B_{if} + g(Z_{2f})'\pi_2 + \delta_1 * A_{if} * Emp_{i\tau f} + \delta_2 * B_{if} * Emp_{i\tau f} + \rho * Emp_{i\tau f} + e_{i\tau f}$$

The estimates of the reform's effects on drop outs, net of the employment response and resulting in means testing, are captured by $\ddot{\beta}_1, \ddot{\beta}_2$ for Type A and B couples, respectively.

[27] We compute the percentage change in employment relative to the percentage change in transfers that each household is eligible for (Table A.1). To ensure that the estimated elasticities are invariant to base specifications, we calculate percentage changes relative to the midpoints between pre and post reform levels. The elasticities for Type A and B couples expressed by the parameters from Equation (2) are:

$$\varepsilon^{Type\ A} = \frac{\beta_1}{2*Start\ Aid}\frac{0.5(2*Start\ Aid+2*SoA)}{(\alpha+0.5\beta_1)}; \qquad \varepsilon^{Type\ B} = \frac{\beta_2}{2*Start\ Aid}\frac{0.5(2*Start\ Aid+2*SoA)}{(\alpha+0.5\beta_2)}$$

where α is the regression constant capturing baseline couples both receiving residency before the reform, and β_1 (β_2) are Type A (B) couples' response relative to baseline couples.

(e.g. -1.78 vs -0.53 in year 1). For males, there is a 60% difference in year 1 (-1.06 vs -0.67), but elasticities are almost identical in year 2.

Panel D in Table 5 shows the household level elasticity of labor earnings with respect to benefit levels. A 1% reduction in benefits increases labor earnings for Type A couples by 1.26% in year 1 and 0.83% in year 2, while the corresponding elasticities for Type B couples are much smaller, at 0.28% and 0.37% for years 1 and 2, respectively.[28] Thus, subtle differences in how transfers are administered post reform between Type A and B couples generate responses that differ by around 200–350%.[29] This finding is in line with Ashenfelter's (1983) result that elasticities depend on the implied tax rates from means testing and non-pecuniary costs of welfare reception, and may help explain how seemingly similar welfare reforms across countries or across time can generate very different labor supply effects.

Is the difference in females' labor force exits in Type B versus Type A couples driven only by females more often being the last arriving spouse (LAS) than males and thus the ones exposed to the disincentive (see discussion above), or by social norms and labor market related gender roles play a role as well? While we have insufficient data on "female first" residencies (males receive residency first in 86% of couples) to separately estimate effects of first versus last residency by gender, Figure A.7 shows labor market outcomes for Type B couples in which the husband receives residency first and the wife second and vice versa by month of LAS residency. The lower unemployment rates after the reform for last residency females is entirely due to higher labor force dropouts, while last residency males also show a weak employment response (Panels A and C of

[28] Table A.5 reports the reform's estimated impact on transfers and labor earnings for the two household types, on the basis of which the elasticities are computed.

[29] While the estimated elasticities of employment and labor earnings with respect to benefits in Table 5 are large (likely reflecting income effects following benefit reductions to a subsistence minimum for a group with high welfare dependency and base labor earnings close to zero), elasticities in this range are not unprecedented. For example, Røed and Zhang (2003) estimate a 0.95 unemployment exit elasticity with respect to the replacement ratio of UI benefits, while Fredriksson and Holmlund (2006) find responses of a benefit reduction on unemployment that correspond to an elasticity as large as -3.

Figure A.7). Thus, while the small samples do not allow us to draw strong conclusions, it seems that the disincentive effects on labor force participation observed for females in Type B couples are also present for last residency males, but to a lesser extent.

4.3.2 Earnings Distributions, and Extensive and Intensive Margin Adjustments

The larger differences between Type A and B couples' labor earnings elasticities at the household level compared with the employment elasticities at the individual level in year 1 (column 4 of Panels C and D in Table 5) could be explained by larger intensive margin response differences within households, or by differences in how the reform affects households' joint responses.

To examine this, we first investigate how the reform changes the distribution of annual labor earnings. We create a series of dummies ($1[y \leq x]$) for whether a household earned \$$x$ or below in the first year following LAS residency, varying x from zero to the top of the earnings distribution. Using these dummy variables as outcomes in estimations of Equation (2), we capture the changes in the cumulative earnings distribution. We graph these distributions for labor earnings on household level in Panel A of Figure 5 and show the differences for Type A and Type B couples relative to the baseline (both receiving residency before the reform) in Panels B and C in Figure 5, respectively. In addition to large extensive margin effects (the differences in distributions at zero earnings), the differences in cumulative earnings distributions for positive earnings levels inform us about potential intensive margin responses. If mass has shifted in the labor earnings distribution at positive earnings levels from pre to post reform, we should observe interline differences in the cumulative distributions that grow when comparing Type A (B) couples with baseline couples (Figure 5B and 5C). The figures show that this is not the case. The interline differences between pre to post reform levels decrease monotonically, which suggests that there are no intensive margin effects of the transfer reduction on labor earnings. Furthermore, the extensive margin responses and the corresponding elasticities (Table 5, Panels C and D) are far larger than the intensive margin elasticities (with respect to the marginal tax rate) previously reported for Denmark (Chetty et al., 2011a; Kleven and Schultz, 2014). Our finding of large extensive margin responses is in line with previous work that finds substantial extensive margin labor supply responses for subpopulations with weak labor market attachment, see e.g. Eissa and Liebman (1996) and Meyer and Rosenbaum (2001).

To investigate whether the differences in the labor earnings distributions between Type A and Type B couples are driven by higher individual income or by joint responses in dual earner households, we define two binary variables. The first is equal to 1 if both spouses find employment in year 1 and zero if one or both do not, and the second is equal to 1 if only one spouse finds employment in year 1 and zero if neither finds employment or both take up a position. Results in Table A.6 (Panel A) along with the corresponding results from Table 5 (shown in Panel B) show that 81% of females' employment response in Type A couples is in dual earner households (6.4 of 7.9 percentage points). Hence, the increased mass at higher labor earnings levels in Type A couples (Figure 5) is induced by dual earner responses. The employment effects in Type B couples, on the other hand, are driven solely by single earner responses where only the male finds employment (a finding that naturally follows from the absence of significant employment effects for females).[30]

4.3.3 Wage Distribution and Reservation Wages

Table 5 and Figure 5 show that the reform induces a total extensive margin response of roughly 20 percentage points for Type A couples but only half of that for Type B couples. Figure A.8 illustrates that the reform creates a further difference in labor supply responses for the two types

[30] Our result that the main distinction between Type A and B couples is secondary earners' (females') labor supply decision supports Eissa and Hoynes' (2004) finding that secondary earner labor supply is very sensitive to the within household incentives following the Earned Income Tax Credit.

of couples when we consider wage rates. In the figure points X, Y1, and Z are identical to X, Y1, and Z in Figure 4. Points X2, Y2, and Z2 represent the tangents of wage rates w_{Pre}, w_A, and w_B at which individuals in the three household types are indifferent between supplying labor and receiving benefits. Unsurprisingly, inducing labor supply responses in the pre reform group, which has the highest benefit level, requires the highest wage rate w_{Pre}. A subtler implication is, however, that the wage rate w_B required for Type B couples to supply labor is higher than w_A for Type A couples. The reason is that Type B couples' outside option changes as one partner can drop out of integration programs without repercussions thereby increasing the non-labor wage rate (see Figure 4).

These predictions are confirmed in the estimated cumulative hourly wage distributions for male spouses (Figure 6), where Type B couples' wages are concentrated at lower levels than pre reform wages, while the lowest wage rates apply to Type A couples.[31] Moreover, estimates in Table A.7 show that only employment in unskilled manual work increases in response to the reform, even for individuals with higher levels of education, which suggests that the skills refugees bring with them have little value in the Danish labor market.[32]

4.4 Long-Run Labor Market Effects

Figure 7 summarizes the estimated long run effects of the reform on the probability of being in employment and not in the labor force (corresponding to Table 3) for adults until 10 years after residency. Labor supply effects were initially considerable in magnitude, close to 10 percentage points on average (Figure 7A) and 15 percentage points for males (Figure 7B). However, after 5-

[31] This distribution is constructed in a similar manner as the previous earnings distribution in Figure 5.
[32] This is in line with Schultz-Nielsen and Skaksen's (2017) finding that non-Western immigrants' returns to education from their origin country are close to zero, and almost all employment is in jobs requiring only compulsory schooling. See Rosholm et al. (2006), Schultz-Nielsen (2008) for discussions of refugee skills in relation to the Danish labor market.

6 years, the reform's employment effects on males are close to zero. Figure 7B shows that the effects on dropping out of the labor force not only are of similar magnitude as the employment effects in the first years after residency, but also more persistent and only become insignificant 7-8 years after residency. Finally, the figure shows that the reform did not promote the longer run integration of refugees into the Danish labor market, as all estimates on labor market outcomes 9-10 years following residency are very close to zero.

5. The Reform's Unintended Consequences

While the lower benefits induce a strong short run employment response, about 70% of households where neither adult takes up employment following the reform experience an income drop by 50% to levels close to a subsistence minimum. This is likely to have consequences along other margins than labor supply, for both adults and their children, which is what we study next.

5.1 Effects on Adult Crime

We first investigate the impact the reduction in benefits has on adults' crime. We report in Table 6 the estimated effects of the reform on the number of crime convictions in years 1 and 4 after residency for all crime, property crime, and theft from supermarkets. Figures A.9 and A.10 present the full set of estimates and pre reform means for both the probability of receiving a crime conviction (A.9) and the number of crime convictions (A.10) and for 1, 2, 3 and 4 years after residency. We present results for all adults, couples, and males and females in couples.

Prior to the reform, refugees' crime rates within their first four years in Denmark were roughly twice as high as average crime rates of native Danes (with equal age distribution) measured within a four-year period, a gap fully explained by refugees' higher levels of property crime. Following the reform, the number of a crime convictions among all refugees (Panel A) increases sharply in the first year after residency (by around 100%), an effect that is driven by property crime (Panel B), namely shoplifting with the main part in supermarkets (Panel C).[33] The table also shows that effects increase until year 4, a pattern driven by males. The reform's effect is even more striking for females. While the pre reform mean is very low, the reform triples the number of crime convictions, an effect that is driven by shoplifting from supermarkets.

Furthermore, extensive and intensive margin responses are very similar, as comparisons of Figures A.9A-C (showing effects on the probability of receiving a crime conviction) and A.10A-C (showing effects on the number of crime convictions) suggest. Thus, the Start Aid reform induced more individuals to commit crimes, rather than increasing simply the number of crimes committed by those who fell criminal also without the reform.

These findings suggest that cutting benefits to or below a subsidence minimum leads to more property crime, even for population groups with very low baseline crime levels such as adult females. This adds individual level data evidence to previous work that identifies effects of day-to-day variation in welfare payments on crime on aggregate level (Foley, 2011; Carr and Packham, 2017), changes to welfare eligibility of criminal offenders and recidivism (Yang, 2017), and state variation in the 1990's U.S. welfare reform implementation and crime rates (Corman et al., 2014).

5.2 Effects on Children

Another important margin along which the large transfer reduction may have effects is on children's lives. Several studies have associated children's well-being with welfare reforms and

[33] Effects on violence are insignificant. There is no evidence of domestic violence.

reductions in their parents' income.[34] We now limit our sample to children and analyze how the reform affects outcomes for three groups who receive residency at different childhood stages: childcare or preschool attendance and language test scores for children receiving residency in early childhood; educational attainment for children receiving residency in middle childhood; and crime for those receiving residency in late childhood.

5.2.1 Childcare and Preschool Attendance, and Language Test Scores

Denmark has two separate preschool institutions: childcare (vuggestue) until age 3 and preschool (børnehave) from age 3 until kindergarten starts at age 6. Both are heavily subsidized.[35] Municipalities are obliged to provide refugee children with childcare and preschool slots from residency and until start of kindergarten. As a visual documentation of the reform's effects on enrollment in childcare or preschool, Panel A of Figure 8 shows enrollment rates for children, who were aged 0-5 at residency, by month of residency. There is a clear discontinuity in enrollment rates around the timing of the reform. The corresponding estimated effects in Panel A of Table 7 show that the likelihood that children attend childcare or preschool during the first two years after residency is 15 percentage points lower post reform. Thus, although low cost universal childcare may improve outcomes for children from low-resource families (Havnes and Mogstad, 2011), and in particular children with migrant backgrounds (Cornelissen et al., 2016), lowering income may keep parents from making use of these services.[36]

[34] See e.g., Aizer et al. (2016), Dahl and Lochner (2012), Duncan et al. (2011), Hoynes et al. (2015), Hoynes et al. (2016), Løken et al. (2012), and Løken et al. 2018 with reviews by Currie and Almond (2011) and Heckman and Mosso (2014).
[35] Childcare and preschool are funded and administered by local municipalities. In our study period, a minimum of 70% of childcare costs were subsidized by the municipality.
[36] Low enrollment rates are seen as a key contributor to the large skill gaps between children with Danish and non-Western backgrounds (Bleses et al., 2016; Eriksen and Hvidtfeldt, 2016). During the past decades, native Danes' enrollment rates have increased towards the present level with universal coverage (97%) for 5 year olds and 85% coverage for 1 year olds, and enrollment rates of children with native background and children with non-Western

One consequence of lower enrollment rates of immigrant children in childcare and preschool could be deficiencies in Danish language proficiency, which has serious consequences for later development and educational opportunities. We therefore study the reform's effect on language test scores for the same children using results from compulsory tests in schools for all children in grades 2 and 4. The test score distributions are graphed by pre and post reform residency in Figure A.12. The figure shows that the test score distributions are centered substantially below the national average (zero on the figure's x-axis) reaffirming the gaps found in earlier comparisons of refugee and native children's test scores (Bleses et al., 2016; Eriksen and Hvidtfeldt, 2016). Moreover, while the figure does not reveal any differences for the modal test scores, a substantial downward shift in test scores for the poorest performing children following the reform is evident. Panel B of Table 7 confirms this finding. Despite no detectable mean effect, the reform results in refugee children being 5 percentage points more likely to earn the lowest test scores from a baseline level of 3%. This is also evident in the reduced-form plot in Panel B of Figure 8, which shows a discontinuity in the probability that children score less than 2.5 standard deviations below the national average.[37] Thus, the reform affects already poor performing children negatively. To understand better what is driving this effect, we estimate the impact of the reform on the likelihood of scoring less than 2.5 standard deviations below the national average separately for the three subsets of the language test. The last three rows in Table 7 show that decoding and reading comprehension are not affected significantly, but the reform leads to significantly worse language comprehension for the least skilled children.

---

background have been gradually converging. The combination of the overall increasing trend and the gradual convergence results in the strong upward trend in enrollment rates in Panel A of Figure 8 and Panel A of Figure A.11.

[37] These findings are in line with those of studies on the association between (or effects of) household income on children's outcomes. For instance, see Hoynes et al. (2015) and Hair et al. (2015) on the association between income (poverty) and infant health / cerebral development, Chetty et al., (2011b), Dahl and Lochner (2012), and Milligan and Stabile (2011) on the effects of EITC / child benefits on test scores, Bastian and Michelmore (2015) on education, Aizer et al. (2016) and Løken et al. (2012) on the effects of income on long-run health, education, and IQ, Akee et al. (2010) on education and crime, and Hoynes et al. (2016) on the effects of food stamps on health.

5.2.2 Long Run Educational Attainment

To assess the reform's long-term impact on children's educational attainment, we use years of completed schooling measured in 2016 for children granted residency at ages 2 to 13 (i.e., 13–15 years after residency was granted between 2001 and 2003).[38] Table 8 reports in the first column the pre reform mean, while the other columns show the mean effects, and effects by gender. The last two rows distinguish between children who were younger or older than 10 years at the time of residency.

While the estimates reveal no statistically significant effects of the reform on average educational attainment, children aged 10–13 at residency complete 0.6 fewer years of schooling, a reduction driven by boys, who complete almost 1 year less schooling because of the reform. The discontinuity in years of completed education for boys aged 10–13 at residency is also illustrated in Panel C of Figure 8. These findings therefore suggest that children in their early teenage years, and particularly boys, are negatively affected by the reform, while children who are younger when receiving residency are not.[39] As a potential mechanism we investigate in Table A.8 how the reform affects the probability that children who receive residency between age 10 and 13 are enrolled in an education, or either in employment or receiving welfare benefits. The table shows that the reform reduces education enrollment for 18-year-olds, who were 10-13 at residency, by 11 percentage points and increases the likelihood that they are either employed or receive welfare accordingly. The effects are, just as the effects on years of completed schooling, most pronounced for boys. It therefore appears that youths substitute away from education to contribute to household subsistence either via youth work or welfare reception.

[38] There is no minimum school leaving age in Denmark but education is compulsory until completion of 9th grade. We use age 2-13 as children younger than 2 at residency had not reached 9th grade by the last year of available data (2016).
[39] Gender differences in responses to reform are in line with Autor et al. (2016) and Chetty et al. (2016a), who find that – measured by educational and labor market outcomes – boys from low income families appear to be particularly vulnerable as they do worse in comparison with girls from similar families.

5.2.3 Crime

As the reform increases adults' crime, it is natural to wonder whether it had a similar effect on adolescents' crime. Providing a visual analysis of the reduced form estimates, Panel D of Figure 8 shows a clear discontinuity around the reform date of convictions for an offense committed at or before age 20 for males aged 14–18 when receiving residency. Panel A of Table 9 shows the overall effects, distinguishing between different crime types. Panel B separates these by gender, and Panel C distinguishes between those who are younger than 14 years versus 14 years or older when receiving residency (i.e. younger vs. older than the age of criminal responsibility in the first year after residency).

The estimates in Panels A and B show that the reform led to significantly more crime convictions overall. Splitting the sample by gender, the table shows an increase in shoplifting for females and both property and violent crime for males.

Breaking estimates down by age group (Panel C) shows that these effects are driven by adolescents who are at least 14 years old when receiving residency. For this age group, the effect of the reform is stark: the number of crime convictions double for basically every crime type, in comparison to the pre reform mean. In sum, we find no evidence of any positive reform effects on refugee children's outcomes but several indications of negative impacts.[40]

[40] To understand what drives the overall effects on adults' and children's crime, and children's daycare enrollment rates, test scores, and educational attainment, we have re-estimated the model conditioning on households' labor supply and how that changes across the reform. Such estimates are likely biased as we condition on a variable (labor earnings) that is itself affected by the reform. With that caveat in mind, estimates suggest that the negative consequences of the reform on adults' crime and children's lives are driven by families who did not respond to the reform by increasing their labor supply, and thus face 40-50% lower disposable incomes following the lower transfer levels.

6 Discussion and Conclusion

The Start Aid reform we evaluate in this paper shares many aspects with reforms recently implemented or currently discussed in other countries. We show that there are indeed large employment responses to the reduction in transfer payments, but these effects are short lived. Moreover, the reform induces large and more persistent female labor force exits, driven by means testing as well as an asymmetric transfer allocation within households, which removes penalization of withdrawal from integration programs and the labor force.

We further illustrate that the vast majority of families affected by the reform do not compensate the substantial drop in transfer income through increased labor supply, and thus experience a large decrease in their disposable income. This leads to a sharp increase in crime, not just for males, but also for females, with female crime being almost exclusively subsistence crime. Perhaps even more concerning than the increase in adults' crime is the large effect the reform has on crime of those who arrive to Denmark as teenagers. Moreover, for younger children the reform not only leads to lower participation rates in preschool programs, but it also lowers performance in language tests as well as the number of years children (and in particular boys) attend school. Overall, the reform affects children negatively in various dimensions, with potentially large negative long-term consequences.

Contrasting the reform-induced increase in labor earnings with the social costs of crime and future earnings losses associated with lower completed education of children, the net effect of the Start Aid reform is likely negative.[41] Thus, while reforms that incentivize labor force participation

[41] We compute the average annual effects per refugee (such that the effects for adults and children are weighted by their respective share in the sample). By this measure the average increase in labor earnings is $493, while the average societal costs due to increasing crime is $397 and $343 for adults and youths, respectively, where crime-costs are monetized using the willingness to pay reported in Cohen and Piquero (2009). The average costs from the lower long-term income of children due to reduced educational attainment, is estimated to be $181 based on the association between income and education for non-Western immigrants who arrived to Denmark as children during

for groups like the one investigated here may have benefits for both taxpayers and target populations, our analysis shows that their success depends on careful implementation of designs that avoids unwanted disincentive effects and ensures persistent employment effects, and accompanying policies to counteract the potential negative consequences, in particular for children.

the 1990s. Therefore, the net effect of the reform amounts to: \$493 - \$397 - \$343 - \$181 = -\$428. Of course, we considered a limited set of outcomes only and there may be negative consequences along other dimensions.

References

Aizer, Anna, Shari Eli, Joseph Ferrie, and Adriana Lleras-Muney. (2016). "The Long-Run Impact of Cash Transfers to Poor Families". *American Economic Review* 106(4): 935-71.

Akee, Randall K. Q., William E. Copeland, Gordon Keeler, Adrian Angold, and E. Jane Costello. (2010). “Parents’ Incomes and Children’s Outcomes: A Quasi-Experiment Using Transfer Payments from Casino Profits”. *American Economic Journal: Applied Economics* 2(1): 86-115.

Andersen, Lars H., Hans Hansen, Marie L. Schultz-Nielsen and Torben Tranæs. (2012). ”Starthjælpens betydning for flygtninges levevilkår og beskæftigelse”. *Rockwool Fondens Forskningsenhed* Arbejdspapir 25. Odense: Syddansk Universitetsforlag.

Ashenfelter, Orley. (1983). “Determining Participation in Income-Tested Social Programs”. *Journal of the American Statistical Association* 78(9): 517-525.

Autor, David, David Figlio, Krzysztof Karbownuk, Jeffrey Roth, and Melanie Wasserman (2016). “Family Disadvantage and the Gender Gap in Behavioral and Educational Outcomes”. NBER Working Paper 22267.

Bastian, Jacob and Katherine Michelmore. (2015). “The Long-Term Impact of the Earned Income Tax Credit on Children’s Education and Employment Outcomes,” *Journal of Labor Economics*, forthcoming.

Bleses, Dorthe, Peter Jensen, Hanne Nielsen, Karen Sehested, and Nina M. Sjö. (2016). *Børns tidlige udvikling og læring*. Copenhagen: Danish Ministry of Children, Education and Gender Equality.

Blundell, Richard and MaCurdy, Thomas, (1999). “Labor supply: A review of alternative approaches”. In O. Ashenfelter and D. Cars (ed.): Handbook of Labor Economics 3, chapter 27, 1559-1695.

Borjas, George J. (2002) “Welfare Reform and Immigration Participation in Welfare Programs”. *The International Migration Review* 36(4) 1093-1123.

Bratberg, Espen, and Kjell Vaage. (2000). “Spell Durations with Long Unemployment Insurance Periods”. *Labour Economics* 7(2): 153-180.

Bratsberg, Bernt, Oddbjørn Raaum, and Knut Røed. (2010). “When Minority Labor Migrants Meet the Welfare State”. *Journal of Labor Economics* 28(3): 633-676.

Brewer, Mike, Marco Francesconi, Paul Gregg, and Jeffrey Grogger. (2009). “In-Work Benefit Reform in a Cross-National Perspective: Introduction.” *Economic Journal* 119(2), 1-14.

Canadian Broadcasting Corporation. (2014). “Omnibus Budget Bill Restrics Refugee Access to Social Assistance”: https://www.cbc.ca/news/politics/omnibus-budget-bill-restricts-refugee-access-to-social-assistance-1.2813994 (accessed 10-12-2018)

Card, David, Raj Chetty, and Andrea Weber. (2007a). “Cash-on-Hand and Competing Models of Intertemporal Behavior: New Evidence from the Labor Market”. *The Quarterly Journal of Economics* 122(4): 1511-1560.

Card, David, Raj Chetty, and Andrea Weber. (2007b). “The Spike at Benefit Exhaustion: Leaving the Unemployment System or Starting af New Job?”. *American Economic Review* 97(2): 113-118.

Carr, Jillian, and Analisa Packham. (2017). “SNAP Benfits and Crime: Evidence from Changing Disbursement Schedules”. Miami University, Department of Economics Working Paper 2017-01.

Chan, Marc K. and Robert A. Moffitt. (2018). “Welfare Reform and the Labor Market”. NBER Working paper no. 24385.

Chetty, Raj, John N. Friedman, Tore Olsen, and Luigi Pistaferri. (2011a). “Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records”. *The Quarterly Journal of Economics* 126(2). 749-804.

Chetty, Raj, John N. Friedman, and Jonah Rockoff. (2011b). “New Evidency of the Long-Term Impacts of Tax Credits”. Washington DC: US Internal Revenue Service.

Chetty, Raj, Nathaniel Hendren, Frina Lin, Jeremy Maherovitz, and Benjamin Scuderi. (2016a). “Childhood Environment and Gender Gaps in Adulthood”. *American Economics Review*: 106(5): 282-288.

Cohen, Mark A. and Alex R. Piquero. (2009). “New Evidence on the Monetary Value of Saving a High Risk Youth.” *Journal of Quantitative Criminology* 25(1): 25-49.

Corman, Hope, Dhaval M. Dave, and Nancy E. Reichman. (2014). “Effects of welfare reform on women's crime”. *International Review of Law and Economics* 40(C): 1-14.

Cornelissen, Thomas, Christian Dustmann, Anna Raute, and Uta Schönberg, (2016) “Who Benefits From Universal Child Care? Estimating Marginal Returns to Early Child Care Attendance,” *Journal of Political Economy*, forthcoming.

Currie, Janet, and Douglas Almond. (2011). “Human Capital Development before Age Five”. In O. Ashenfelter, and D. Card (ed.): *Handbook of Labor Economics* 4(B), chapter 15. 1315-1486.

Dahl, Gordon B, and Lance Lochner. (2012).”The Impact of Family Income on Child Achievement: Evidence from the Earned Income Tax Credit“. *American Economic Review* 102(5): 1927-956.

Danish Prime Minister’s Office. (2002). *På vej mod en ny integrationspolitik*. Copenhagen: Danish Prime Minister’s Office: 1-20.

Duncan, Greg J., Pamela A. Morris, and Chris Rodrigues. (2011). "Does Money Really Matter? Estimating Impacts of Family Income on Young Children's Achievement with Data from Random-Assignment Experiments". *Developmental Psychology* 47(5): 1263-1279.

Eissa, Nada, and Jeffrey Liebman. (1996).”Labor Supply Response to the Earned Income Tax Credit“. *The Quarterly Journal of Economics* 111(2): 605-637.

Eissa, Nada, and Hilary Williamson Hoynes. (2004).”Taxes and the Labor Market Participation of Married Couples: The Earned Income Tax Credit“. *Journal of Public Economics* 88(9-10): 1931-1958.

Eriksen, Hanne-Lise F., and Camilla Hvidtfeldt. (2016). Indskolingselevers Trivsel og Faglige Kompetencer, Resultater fra Høje-Taastrup Kommune. Odense: Syddansk Universitetsforlag.

Falk, Gene. (2014). “Temporary Assistance for Needy Families (TANF): Eligibility and Benefit Amounts in State TANF Cash Assistance Programs”. Congresional Research Service, R43634.

Foley, C. Fritz. (2011). “Welfare Payments and Crime”. *The Review of Economics and Statistics* 93(1): 97-112.

Frankfurter Allgemein. (2016). “Ohne Integration Werden die Leistungen Gekürzt” [Social Assistance Should be Shortened in the Absence of Integration]: https://www.faz.net/aktuell/politik/inland/andrea-nahles-fordert-fluechtlinge-auf-sich-zu-integrieren-14044777.html (accessed 10-12-2018).

Fredriksson, Peter, and Bertil Holmlund. (2006). “Improving Incentives in Unemployment Insurance: A Review of Recent Research”. *Journal of Economic Surveys* (20)3, 357-386.

Hair, Nicole, Jamie L. Hanson, Barbara L. Wolfe, and Seth D. Pollak, (2015). “Association of Child Poverty, Brain Development, and Academic Achivement”. *JAMA Pediatric* 169(9). 822-829.

Havnes, Tarjei, and Magne Mogstad. (2011). "No Child Left Behind: Subsidized Child Care and Children's Long-Run Outcomes". *American Economic Journal: Economic Policy* 3(2): 97-129.

Heckman, James J. (1993). “What Has Been Learned about Labor Supply in the Past Twenty Years?”. *American Economic Review* 83(2), 116-121.

Heckman, James J., and Stefano Mosso. (2014). “The Economics of Human Development and Social Mobility”. *Annual Review of Economics* 6(1): 689-733.

Hoff, Erika, Brett Laursen, and Twila Tardif. (2002). “Socioeconomic Status and Parenting”. In M. H. Bornstein (ed.) *Handbook of Parenting, Volume 2, Biology and Ecology of Parenting*, New Jersey: Lawrence Erlbaum Associates, Publishers: 231-252.

Hoynes, Hillary Williamson. (1996). “Welfare Transfers in Two-Parent Families: Labor Supply and Welfare Participation Under AFDC-UP”. *Econometrica* 64(2): 295-332.

Hoynes, Hillary Williamson, Douglas L. Miller, and David Simon. (2015). “Income, the Earned Income Tax Credit, and Infant Health”. *American Economic Journal: Economic Policy* 7(1): 172-211.

Hoynes, Hilary Williamson, Diane Whitmore Schanzenbach, and Douglas Almond. (2016). ”Long Run Impacts of Childhood Access to the Safety Net”. *American Economic Review* 106(4): 903-934.

Huynh, Duy T., Marie L. Schultz-Nielsen, and Torben Tranæs. (2007). “Employment Effects of Reducing Welfare to Refugees. *Rockwool Foundation Research Unit* Study Paper 15.

Hvidtfeldt, Camilla, Marie L. Schultz-Nielsen, Erdal Tekin, and Mogens Fosgerau. (2017). “Asylum Process and Employment among Refugees: How Bad is Waiting Time?”. Working paper.

Immervoll, Herwig, Henrik Jakobsen Kleven, Claus Thustrup Kreiner, and Emmanuel Saez. (2007). “Welfare Reform in European Countries: A Microsimulation Study,” *The Economic Journal* 117, 1-44.

Kleven, Henrik Jakobsen, and Esben Schultz, (2014). ”Estimating Taxable Income Responses Using Danish Tax Reforms”. *American Economic Journal: Economic Policy* 8(4). 271-301.

Krueger, Alan B. and Bruce D. Meyer. (2002). “Labor Supply Effect of Social Insurance”. *Handbook of Public Economics* 4: 2327-92.

Kyyra, Tomi and Virvi Ollikainen. (2008). “To Search or Not to Search? The Effects of UI Benefits Extension on Older Unemployed”. *Journal of Public Economics* 92(10-11) : 2048-2070.

Kilström, Mathilda, Birthe Larsen, and Elisabet Olme. 2018. “Should I stay or Must I Go? Temporary Protection and Refugee Outcomes”. Unpublished working paper.

Lemiuex, Thomas and Kevin Milligan. (2008). “Incentive Effects of Social Assistance: A Regression Discontinuity Approch”. *Journal of Econometrics* 142(2): 807-828.

Library of Congress: http://www.loc.gov/law/foreign-news/article/germany-act-to-integrate-refugees-enters-into-force/ (accessed 10-12-2018)

Løken, Katrine V., Magne Mogstad, and Matthew Wiswall. (2012).”What Linear Estimators Miss: The Effects of Family Income on Child Outcomes”. *American Economic Journal: Applied Economics* 4(2): 1-3.

Løken, Katrine, Kjell Erik Lommerud, Katrine Holm Reinso. (2018). “Single Mothers and their Children: Evaluating a Work-Encouraging Welfare Reform”. *Journal of Public Economics* 167(11): 1-20.

Matthiessen, Poul C. (2009). *Immigration to Denmark. An Overview of Research Carried Out from 1999 to 2006 by the Rockwool Foundation Research Unit*. Copenhagen: The Rockwool Foundation Research Unit and University Press of Southern Denmark.

Meyer, Bruce D., and Dan T. Rosenbaum. (2001). ”Welfare, the Earned Income Tax Credit, and the Labor Supply of Single Mothers“. *The Quarterly Journal of Economics* 116(3): 1063-114.

Milligan, Kevin, and Mark Stabile. (2011). "Do Child Tax Benefits Affect the Well-being of Children? Evidence from Canadian Child Benefit Expansions". *American Economic Journal: Economic Policy* 3(3): 175-205.

Mogstad, Magne and Chiara Pronzato. (2012). “Are Lone Mothers Responsive to Policy Changes? Evidence from a Workfare Reform in a Generous Welfare State”. *The Scandinavian Journal of Economics* 114(4): 1129-1159.

Moffitt, Robert A. (2002). ”Economic Effects of Means-Tested Transfers in the U.S.” In Poterba, James M. (ed.). *Tax Policy and the Economy* 16.: 1-36.

Moffitt, Robert A. (ed.). (2015). *Economics of Means-Tested Transfer Programs in the United States*. Chicago: National Bureau of Economic Research.

OECD Migration Outlook 2006-2018: https://www.oecd-ilibrary.org/social-issues-migration-health/international-migration-outlook-2018_migr_outlook-2018-en (accessed 10-12-2018)

OECD Trends in International Migration 1997-2004: https://www.oecd-ilibrary.org/social-issues-migration-health/trends-in-international-migration_20746873 (accessed 10-12-2018)

Pedersen, Peder J. (2010) “Immigration and Welfare State Cash Benefits: The Danish Case”, IZA Discussion Paper no. 6220.

Regierungsprogramm. (2017). “Zusammen. Für unser Österreich” [Together. For our Austria]: https://www.dieneuevolkspartei.at/download/Regierungsprogramm.pdf (accessed 10-12-2018).

Rosholm, Michael, and Rune Vejlin. (2010). "Reducing Income Transfers to Refugee Immigrants: Does Starthelp Help You Start?". *Labour Economics* 17(1): 258-275.

Rosholm, Michael, Kirk Scott, and Leif Husted. (2006). "The Times They Are A-Changin': Organizational Change and Immigrant Employment Opportunities in Scandinavia". *International Migration Review* 40(2): 318-347.

Røed, Knut, and Tao Zhang. (2003). "Does Unemployment Compensation Affect Unemployment Duration?". *The Economic Journal* 113(484): 190-206.

Saez, Emmanuel, (2002). “Optimal Income Transfer Programs: Intensive versus Extensive Labor Supply Responses”. *The Quarterly Journal of Economics* 117(3). 1039-1073.

Saez, Emmanuel, Joel Slemrod, and Seth H. Giertz, (2012). “The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review”. *Journal of Economic Literature* 50(1). 3-50.

Schultz-Nielsen, Marie L. (2008). ”Økonomisk Afkast af Uddannelse”. In T. Tranæs (ed.): *Indvandrerne og det danske uddannelsessystem*. København: Rockwool Fondens Forskningsenhed og Gyldendal: 193-211.

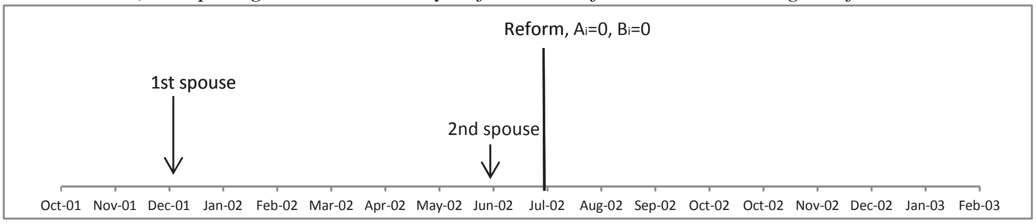
Schultz-Nielsen, Marie L. and Jan R. Skaksen. (2017). ”Indvandreres uddannelse”, Rockwool Fondens Forskningsenhed Study Paper 48.

Swissinfo.ch. (2017). ”Zurich Cuts Funding for Temporary Asylum Seekers”: https://www.swissinfo.ch/eng/unwanted_zurich-cuts-funding-for-temporary-asylum-seekers/43544010 (accessed 10-12-2018).

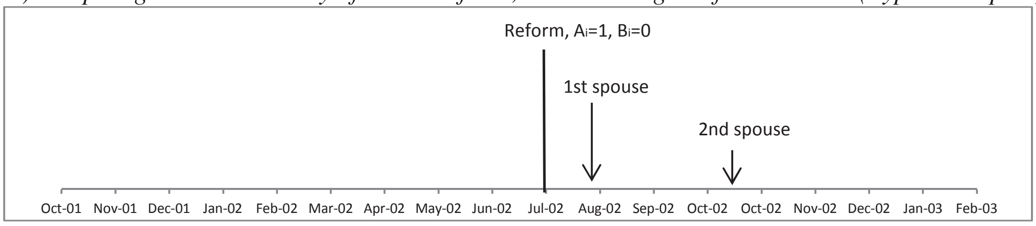
Yang, Crystal S., (2017) “Does Public Assistance Reduce Recidivism*?” American Economic Review: Papers and Proceedings* 107(5): 551-555.

Figure 1. Treament status of couples depending on timing of residency.

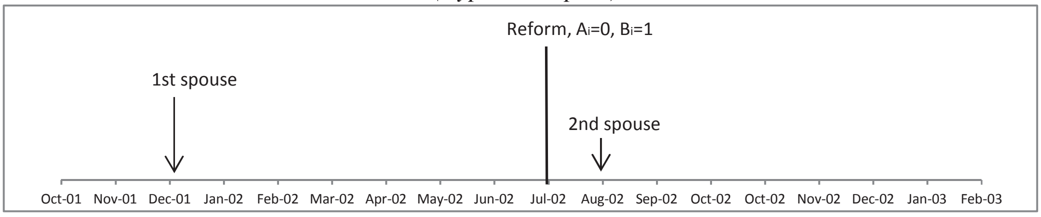
A) Couples granted residency before the reform; both are eligible for SoA



B) Couples granted residency after the reform; both are eligible for Start Aid (Type A couples)



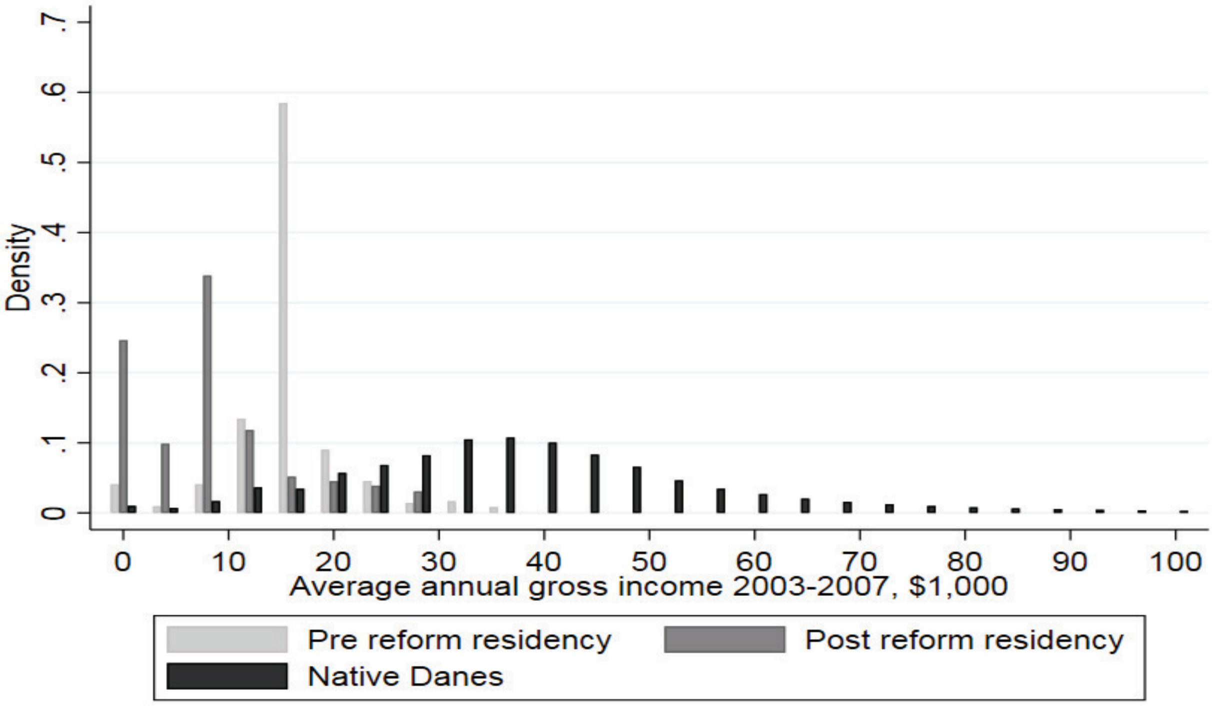
C) Couples granted residency on both sides of the reform; 1st spouse eligible for SoA, 2nd ineligible for benefits (Type B couples)



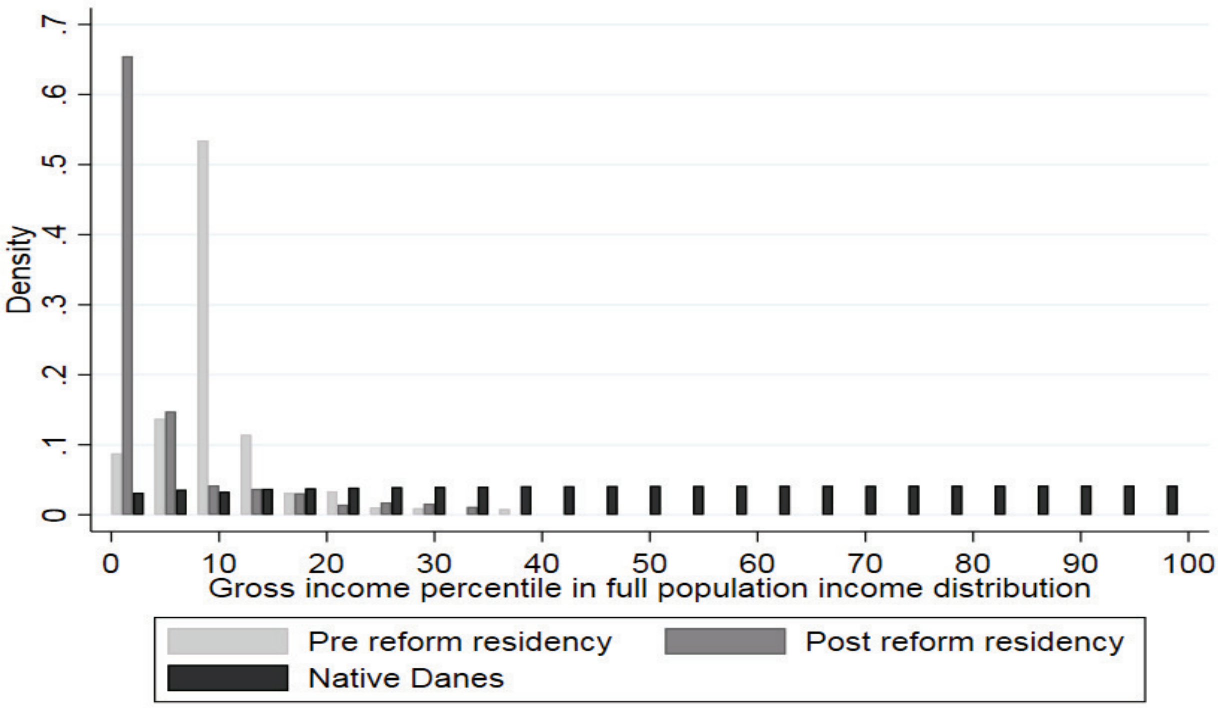
Note: The figure shows the three different treatment states depending on the timing of each spouse's residency. Panel A: Both receive residency before the reform and are thus eligible for full pre-reform SoA. Here both of the treatment dummy variables Ai and Bi = 0. Panel B (with the treatment dummy Ai = 1): *Type A couples* where both spouses receive residency after the reform and are thus eligible for Start Aid. Panel C (with the treatment dummy Bi = 1): *Type B couples* where one spouse receives residency before the reform and one after. The treatment dummies Ai and Bi are detailed in Section 3.

Figure 2. Income distributions for refugees by pre and post reform residency, and native Danes.

A) Distribution of absolute income levels

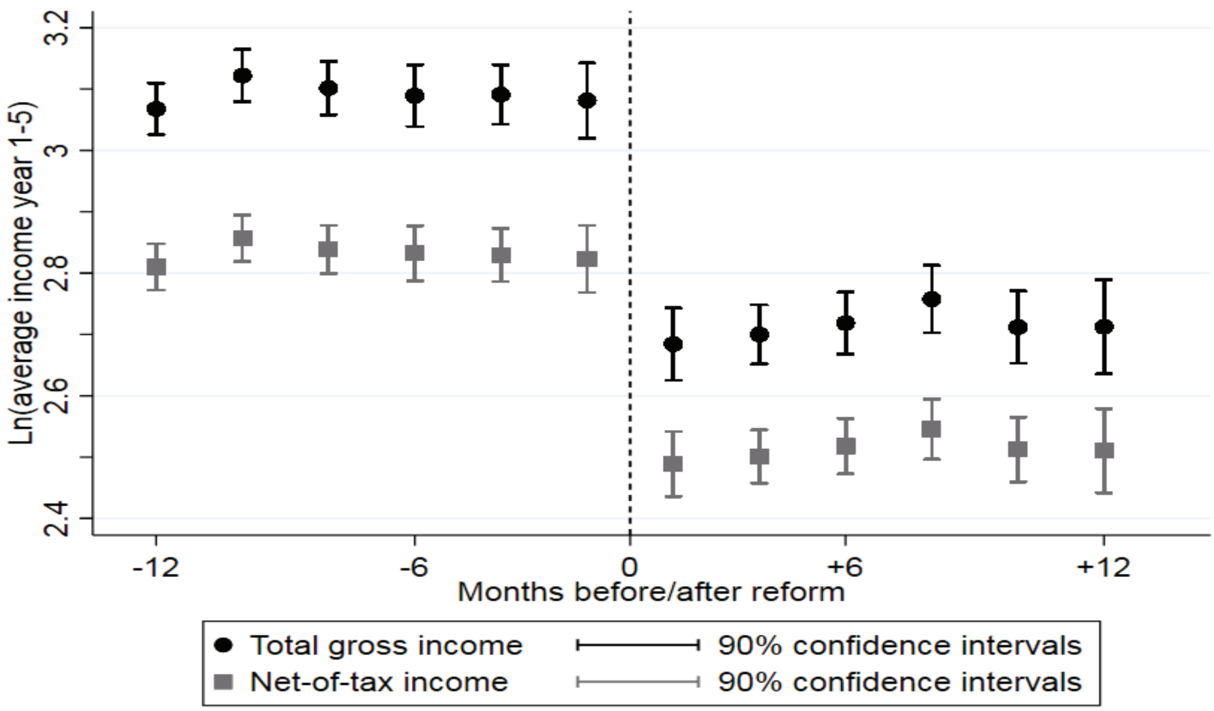


B) Distribution of income percentiles in the full population income distribution



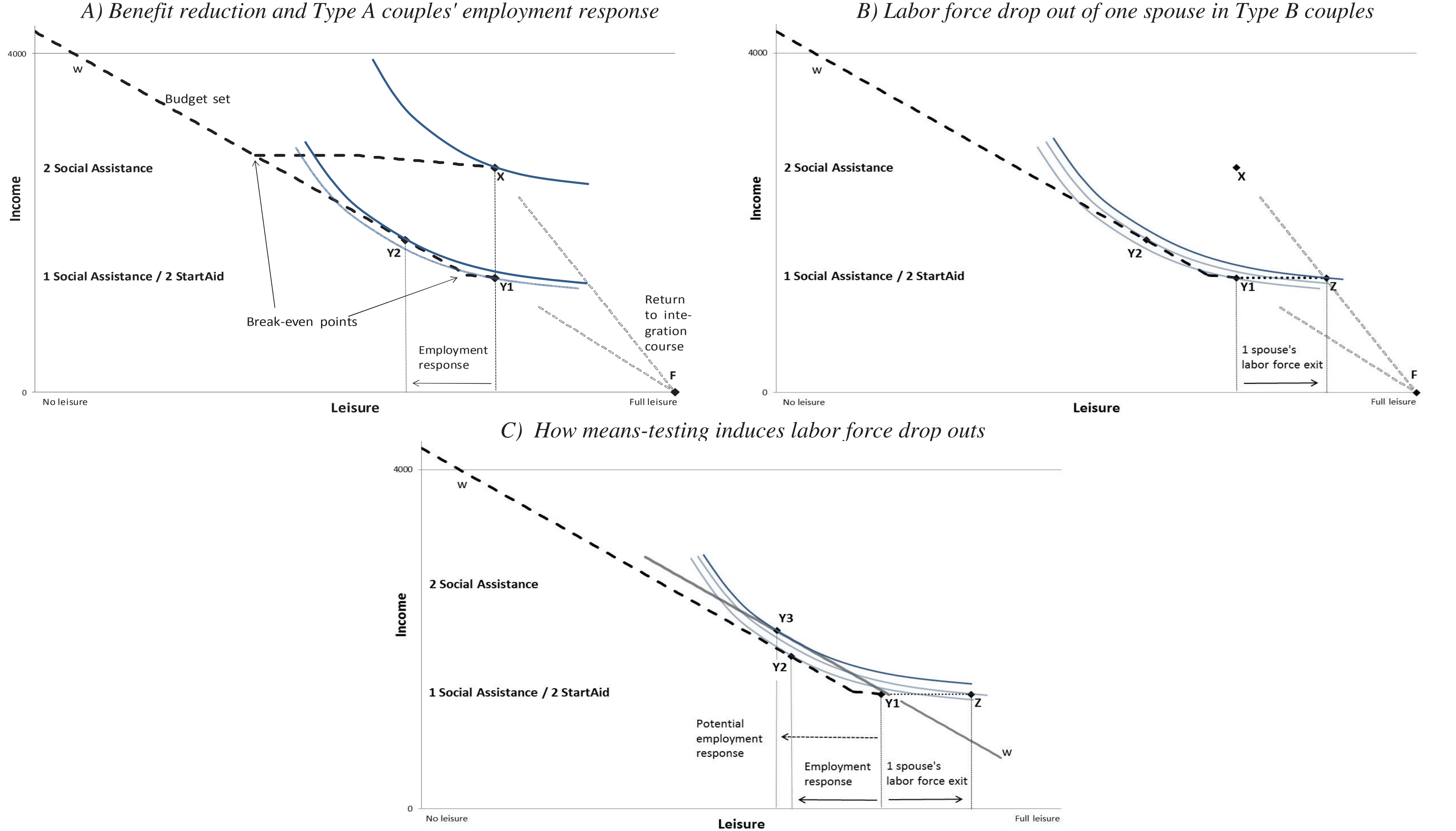
Note: The figure shows the income distributions of adult refugees (age 30 or above) receiving residency in 2002 by whether they received residency before the reform (eligible for SoA) or after the reform (eligible for Start Aid) and adult native Danes (age 30 or above) for comparison. The income distribution is based on total gross income from 2003-2007. Panel A presents the distribution of income levels and Panel B presents, for each of the three groups, the distribution of income percentiles (in the full population income distribution).

Figure 3. Ln(total individual income), the first five years after residency.



The figure shows average and 90% confidence intervals of log of total gross income, and net-of-tax income in year 1-5 after residency measured on individual level for the base sample of adults (age 18-55 on the time of residency) by bi-monthly bins of residency timing.

Figure 4. How the reform affects household budget constraints and labor supply.



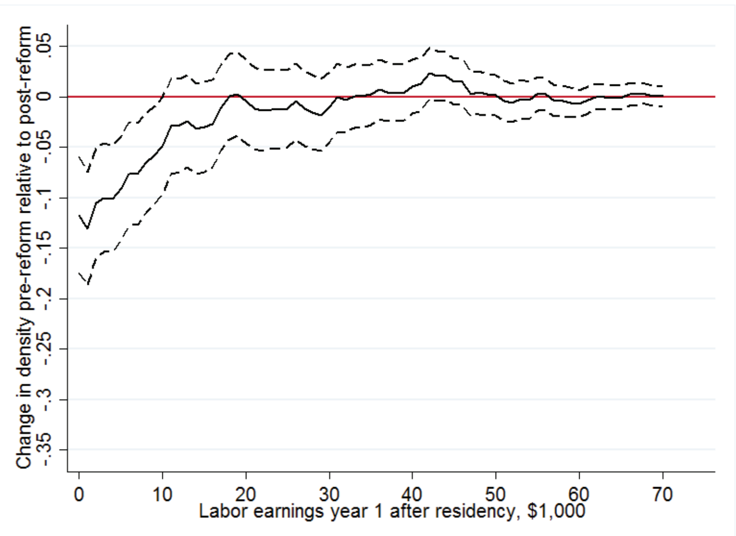
Note: Figure treats households as unitary and assumes convex preferences. Panel A: Reform reduces couples' benefits X→Y1 (conditional on program participation) resulting in lower return to program participation (the lower slope of Y1-F relative to X-F), and Y1→Y2 illustrates Type A couples' employment response. Panel B: For Type B couples it is optimal to let one spouse drop out of the labor force and increase leisure: Y1→Z. Panel C: The dashed line with slope w is the budget set had there been no means testing. Y3 is the potential employment response dominating points Z, Y2.

Figure 5. Change in labor earnings distribution as result of the Start Aid reform, by household type.

A) Change in cumulated yearly labor earnings distribution

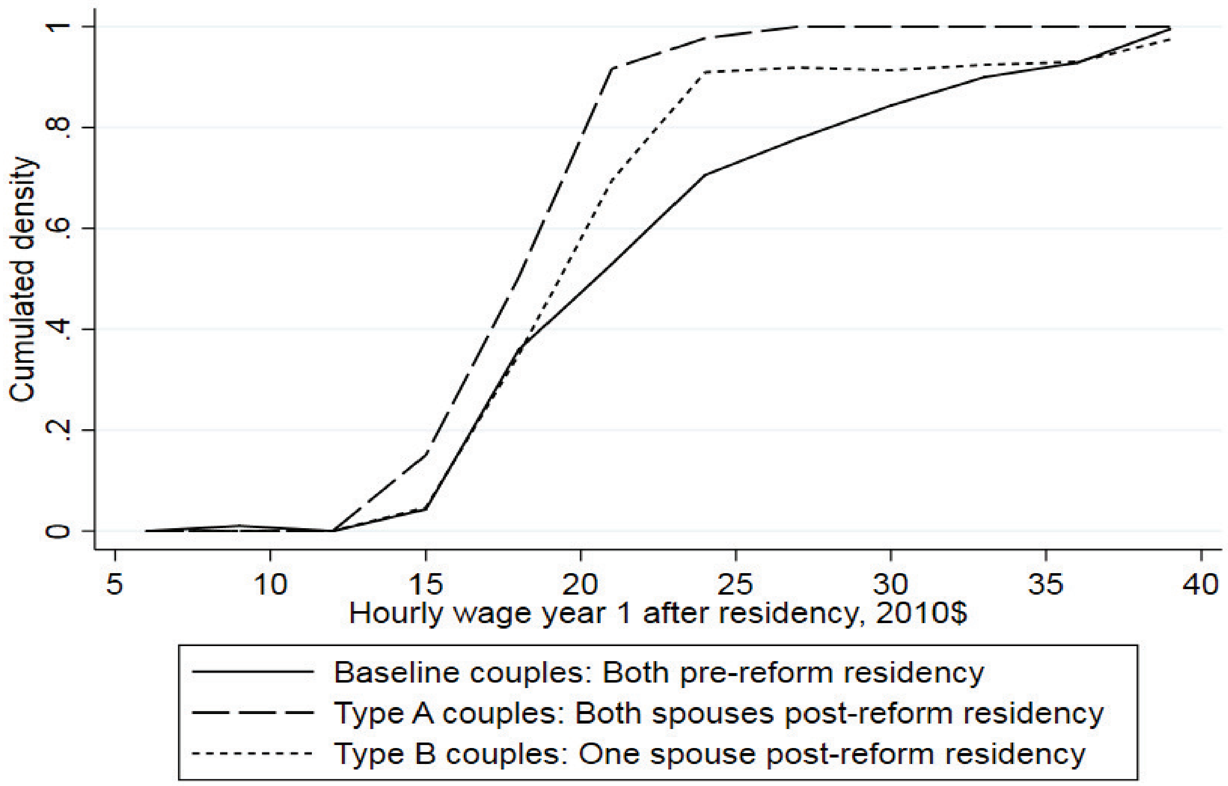
B) Change in density in yearly labor earnings distribution, Type A couples relative to baseline

C) Change in density in yearly labor earnings distribution, Type B couples relative to baseline



Note: The figure shows the estimated changes in cumulative annual labor earnings distributions for couples on household level in Panel A. Panels B and C show the changes induced by the reform (the vertical differences between the lines in Panel A) for couples where both receive recidency after the reform and for couples where only one spouse does, along with the corresponding 95% confidence intervals.

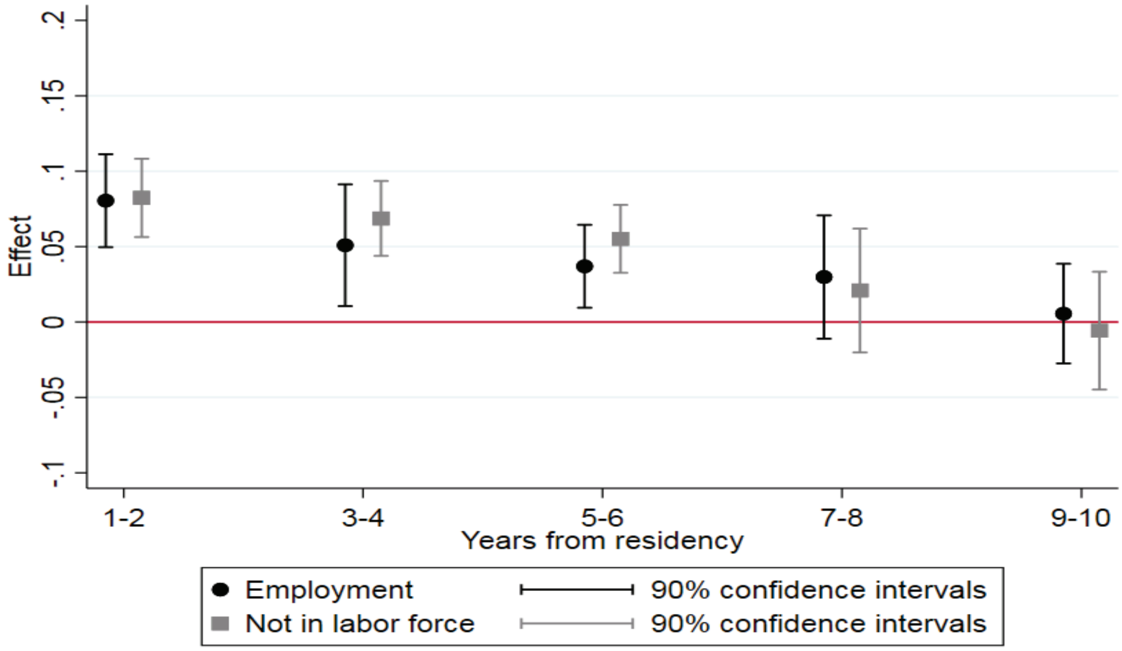
Figure 6. Change in distributions of hourly wage rates as result of the Start Aid reform, by household type, males in couples, year 1.



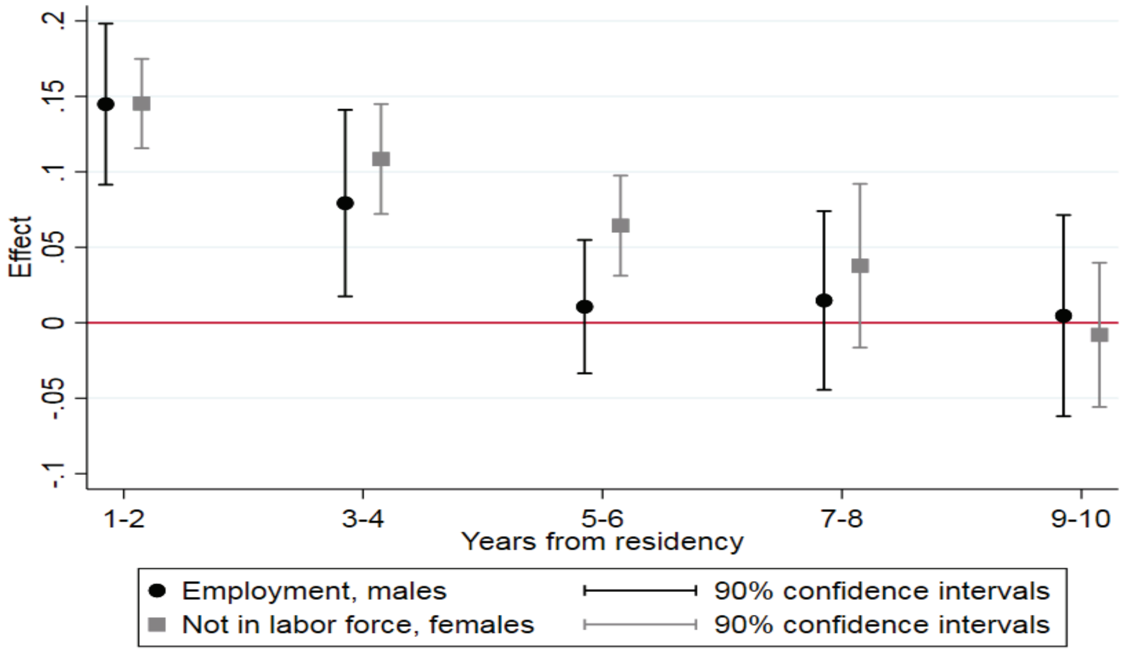
Note: The figure shows the distributions of hourly wage rates. The figure shows the estimated effects using Equation (2) and as outcomes dummy variables indicating a wage rate <= each level from 0 to 40.

Figure 7. Effect of Start Aid reform on employment and not in the labor force (NILF) rates, year 1-10 after residency.

A) Effects for all adults

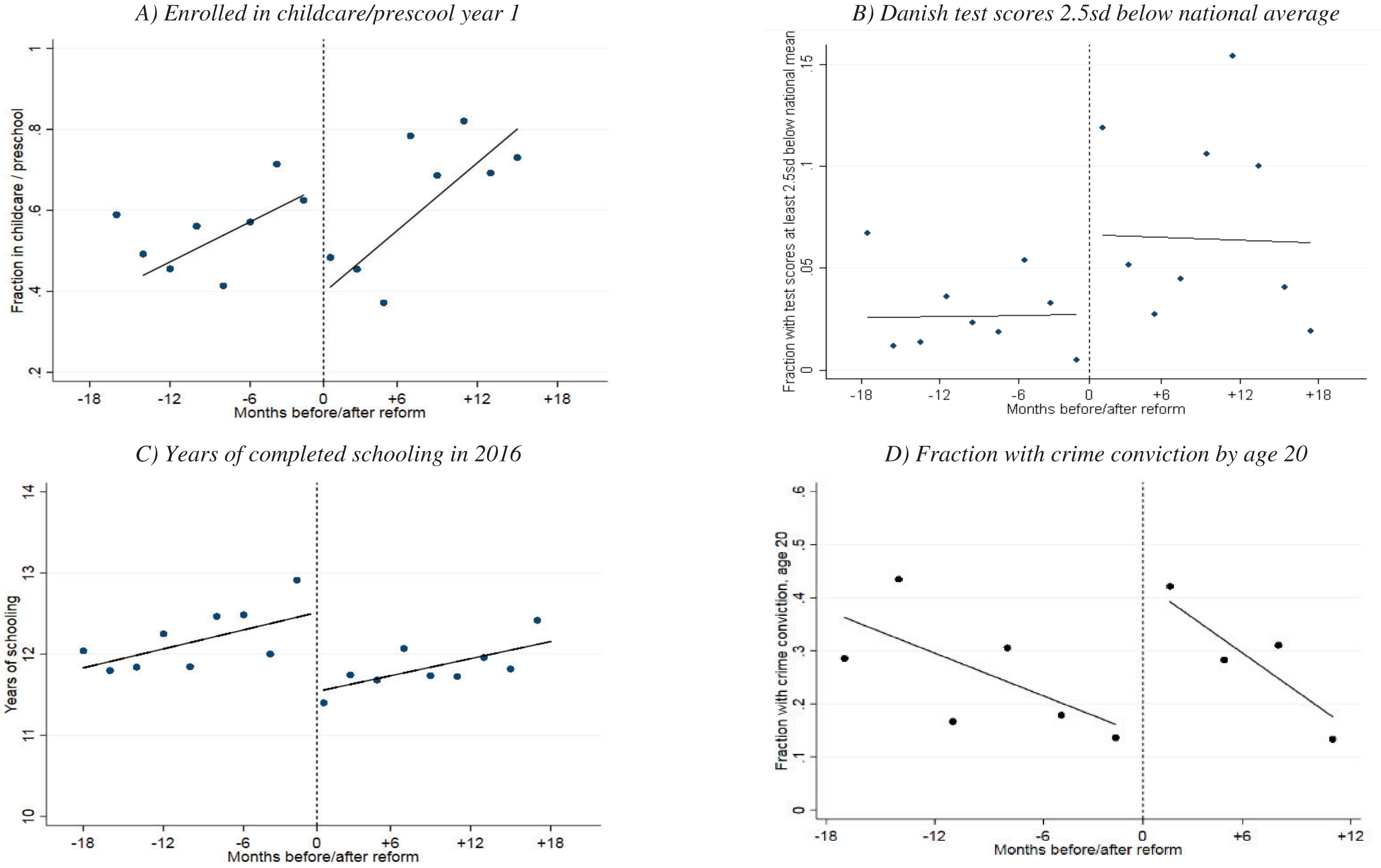


B) Effects by gender



Note: The figure shows estimated effects of being granted residency after the reform relative to before the reform on the subsequent probability of being employed or not in the labor force measured for the base sample of adults (age 18-55 on the time of residency and younger than 60 at the time of measurement) in year 1-10 after residency (biannual averages). Panel A shows the estimates for the full sample and Panel B shows the employment estimates for males and not in labor force estimates for females. All estimates are conditional on the running variable, covariates (see Table 2), and year fixed effects.

Figure 8. Reduced form figures of secondary outcomes, binned by residency relative to the reform.



Note: The figures show childcare/preschool enrollment and Danish test scores for children age 0-5 at time of residency (Panels A-B), years of completed schooling for males aged 10-13 at time of residency (Panel C), and fraction of males aged 14-18 with a crime conviction by age 20 (Panel D) by monthly bins of timing of residency. The dashed vertical line indicates the timing of the reform in July 2002. Statistics Denmark require that figures must contain more than 5 observations per cell. Low frequency outcomes such as in Panel D cut for >+12 months when the low frequency cells appear.

Table 1. Sample means of covariates, base sample aged 18-55 at residency and couples sample.

	A) Base sample			*B) Couples sample*				
	All	Pre reform residency	Post reform residency	All	Pre reform residency	Post reform residency	Males	Females
Reform=1	0.371	0.000	1.000	0.270	0.000	1.000	0.163	0.378
	(0.483)	(0.000)	(0.000)	(0.444)	(0.000)	(0.000)	(0.369)	(0.485)
Age at residency	32.625	32.700	32.498	33.294	33.238	33.446	35.403	31.186
	(8.270)	(8.311)	(8.202)	(8.538)	(8.775)	(7.862)	(8.556)	(7.983)
1 or more siblings	0.017	0.014	0.022	0.002	0.003	0.000	0.001	0.002
	(0.129)	(0.118)	(0.146)	(0.044)	(0.052)	(0.000)	(0.038)	(0.050)
Female	0.507	0.475	0.560	0.500	0.426	0.699	0.000	1.000
	(0.500)	(0.499)	(0.496)	(0.500)	(0.495)	(0.459)	(0.000)	(0.000)
# of children	2.257	2.346	2.106	2.411	2.414	2.405	2.411	2.411
	(1.903)	(1.931)	(1.847)	(1.265)	(1.278)	(1.232)	(1.265)	(1.265)
Single	0.246	0.226	0.279	-	-	-	-	-
	(0.431)	(0.418)	(0.449)					
Muslim countries	0.838	0.878	0.769	0.843	0.859	0.799	0.844	0.841
	(0.369)	(0.327)	(0.422)	(0.364)	(0.348)	(0.401)	(0.363)	(0.365)
Iraq	0.520	0.582	0.416	0.546	0.562	0.505	0.554	0.539
	(0.500)	(0.493)	(0.493)	(0.498)	(0.496)	(0.500)	(0.497)	(0.499)
Iran	0.071	0.045	0.115	0.061	0.050	0.092	0.060	0.062
	(0.257)	(0.208)	(0.319)	(0.240)	(0.218)	(0.289)	(0.237)	(0.242)
Eastern Europe/former USSR	0.055	0.050	0.062	0.042	0.035	0.062	0.034	0.050
	(0.227)	(0.219)	(0.241)	(0.201)	(0.184)	(0.241)	(0.182)	(0.218)
Rest of the world	0.108	0.071	0.170	0.115	0.106	0.139	0.121	0.109
	(0.310)	(0.257)	(0.375)	(0.319)	(0.308)	(0.346)	(0.327)	(0.311)
Refugee permit status	0.618	0.635	0.588	0.527	0.590	0.387	0.847	0.257
	(0.486)	(0.482)	(0.492)	(0.499)	(0.492)	(0.487)	(0.360)	(0.437)
First arriving spouse in couple	-	-	-	0.500	0.601	0.228	0.862	0.138
				(0.500)	(0.490)	(0.420)	(0.344)	(0.344)
Observations	4,843	3,044	1,799	4,072	2,972	1,100	2,036	2,036

Note: The table shows sample means (with standard deviations in parentheses) for the base sample of adults receiving recidency +/- 18 months around the reform and the balanced couples sample of all couples where at least one spouse receives residency +/-18 months around the reform. '# of children' refers to the number of children upon residency. 'Muslim countries' refers to countries where the majority are muslim and not the individual's religion. 'Refugee' refers to the individual receving recidency based on refugee status (is residency given on grounds of being a refugee 1, or from being the spouse of an individual with refugee status, 0). 'First arriving spouse' likewise refers to being the first spouse in the couple who is granted residency in Denmark. A minor share of spouses do not have the same country of origin. This is almost exclusively couples with origin in two neighboring countries (with loosely enforced borders, low population density) as for example Somalia and Eritrea. 'Single' is left out of the couples sample summary as no one is single by construction. First arriving spouse is left out of the base sample summary, as this is only relevant in the couples sample.

Table 2. Conditional balancing test of covariates across reform.

	A) Base sample	*B) Full sample*	*C) Couples sample*
Age at residency	-0.000	-0.000	-0.002**
	(0.001)	(0.001)	(0.001)
Any siblings	0.007	-0.007	-
	(0.030)	(0.022)	
Female	0.015	0.005	0.014
	(0.010)	(0.006)	(0.024)
# of children	-0.007	-	-0.010*
	(0.004)		(0.006)
Single	0.007	-	-
	(0.018)		
Spouse present in Denmark	0.016	-	-
	(0.016)		
Eastern Europe/former USSR	-0.036	(0.027)	-0.043
	(0.028)	-0.015	(0.038)
Rest of the world	-0.013	(0.023)	0.011
	(0.025)	-0.001	(0.041)
Refugee permit status	-0.003	(0.016)	0.023
	(0.018)	(0.057)	(0.037)
First residency in couple	-	-	-0.008
			(0.030)
F	1.010	0.545	1.145
P(F)	0.451	0.794	0.359
Observations	4,843	8,142	4,072
Running variable	X	X	X

Note: Panels A and B show full regression results and F-tests of conditional balancing of covariates across the reform. Panels A and B show results from regressing a dummy indicating whether residency was granted before or after the reform on all covariates and the running variable for the main sample (age 18-55) and the full sample (including children). Panel C shows the equivalent results for the couples sample including a dummy indicating whether the spouse is granted residency before or after the reform. The table hence reports the individual γ's and an F-test for joint significance of the γ's from the regression:

$$reform(0/1) = X'\gamma + f(Z)'\pi$$

with standard errors in parentheses. Covariates include age at residency, whether the refugee has any siblings (except for couples in Panel C due to lack of variation in siblings in the spouse sample), gender, number of children (except for Panel B as this sample includes children), marital status (except for Panel B as this sample includes children), spouse present in Denmark (except for Panel B as this sample includes children), country of origin (Eastern Europe/former USSR and rest of world, with predominantly muslim countries as reference category), and refugee permit status (is residency given on grounds of being a refugee 1, or from being the spouse / child of an individual with refugee status, 0). The covariates for Panel C also include a dummy for whether spouse is the first or last to receive residency.

* p<0.1; ** p<0.05; *** p<0.01

Table 3. Effect of reform on subsequent annual individual transfers, labor earnings (both measured in USD 1,000), employment, unemployment, and fraction not in the labor force.

	A) Transfers		*B) Labor earnings*		*C) Employment rate*		*D) Unemployment rate*		*E) Not in labor force*	
Years since residency	Pre reform mean	Estimate	Pre reform mean	Estimate	Pre reform mean	Estimate	Pre reform mean	Estimate	Pre reform mean	Estimate
1	18.431	-9.775***	1.852	1.144***	0.103	0.092***	0.868	-0.164***	0.028	0.072***
	(7.663)	(0.407)	(6.353)	(0.400)	(0.305)	(0.022)	(0.339)	(0.027)	(0.164)	(0.014)
2	17.979	-8.320***	4.182	1.567***	0.188	0.070***	0.755	-0.158***	0.053	0.093***
	(8.655)	(0.446)	(10.401)	(0.541)	(0.391)	(0.019)	(0.430)	(0.028)	(0.224)	(0.020)
3-5	15.849	-4.956***	8.424	1.070**	0.323	0.041*	0.581	-0.104***	0.093	0.066***
	(8.760)	(0.457)	(13.273)	(0.451)	(0.390)	(0.021)	(0.401)	(0.015)	(0.230)	(0.012)
Observations	3,044	4,843	3,044	4,843	3,044	4,843	3,044	4,843	3,044	4,843

Note: The table shows the estimated effects of being granted residency after the reform relative to before the reform on subsequent transfer income and labor earnings on individual level, and the probability of being employed, unemployed, or not in the labor force measured for the base sample of adults (age 18-55 on the time of residency) in year 1, 2, and the average of years 3-5 since residency. The table also shows pre reform means of the outcome variables. All estimates are conditional on the running variable, covariates (see Table 2), and year fixed effects.

* p<0.1; ** p<0.05; *** p<0.01

Table 4. Effect of reform on subsequent labor market outcomes, by gender.

Years since residency	*A) Employment rate*		*B) Unemployment rate*		*C) Not in labor force*		*D) Dropped out*		*E) Disability benefits*	
	Males	Females	Males	Females	Males	Females	Males	Females	Males	Females
1	0.160***	0.037	-0.155***	-0.171***	-0.004	0.132***	0.002	0.126***	-0.006	0.007
	(0.044)	(0.022)	(0.045)	(0.022)	(0.012)	(0.021)	(0.010)	(0.021)	(0.007)	(0.009)
2	0.132***	0.015	-0.137***	-0.172***	0.011	0.162***	0.020	0.159***	-0.010	0.003
	(0.035)	(0.019)	(0.038)	(0.029)	(0.023)	(0.025)	(0.017)	(0.025)	(0.013)	(0.005)
3-5	0.042	0.041**	-0.069***	-0.130***	0.029	0.093***	0.013	0.096***	0.016	-0.003
	(0.033)	(0.016)	(0.024)	(0.024)	(0.024)	(0.019)	(0.012)	(0.018)	(0.017)	(0.010)
Observations	2,390	2,453	2,390	2,453	2,390	2,453	2,390	2,453	2,390	2,453

Note: The table shows the estimated effects, by gender, of being granted residency after the reform relative to before the reform on subsequent probability of being employed, unemployed, not in the labor force, dropped out of the welfare system completely, or on disability benefits for the base sample of adults (age 18-55 on the time of residency) in year 1 and 2, and the average of years 3-5 since residency. All estimates are conditional on the running variable, covariates (see Table 2), and year fixed effects.

* p<0.1; ** p<0.05; *** p<0.01

Table 5. Effect of reform on subsequent labor market outcomes, by gender and household type.

Treatment	Year since residency	(1) Employment	(2) Unemployment	(3) NILF	(4) Dropped out	(5) Dropped out, corrected for means testing
A) Type A Couples, both granted residency after reform						
Males	1	0.153***	-0.202***	0.049	-0.006	-0.013
		(0.043)	(0.050)	(0.037)	(0.026)	(0.009)
	2	0.095*	-0.134**	0.039	0.011	0.004
		(0.050)	(0.058)	(0.046)	(0.030)	(0.011)
Females	1	0.079*	-0.169***	0.090**	0.084***	-0.030
		(0.043)	(0.050)	(0.037)	(0.026)	(0.019)
	2	0.081	-0.110*	0.031	0.058*	0.018
		(0.050)	(0.058)	(0.046)	(0.030)	(0.022)
B) Type B Couples, one granted residency after reform						
Males	1	0.075**	-0.101***	0.025	0.020	0.012**
		(0.031)	(0.036)	(0.027)	(0.019)	(0.006)
	2	0.107***	-0.137***	0.029	0.026	0.013
		(0.036)	(0.042)	(0.033)	(0.022)	(0.008)
Females	1	0.031	-0.204***	0.173***	0.154***	0.096***
		(0.031)	(0.036)	(0.027)	(0.019)	(0.019)
	2	0.036	-0.209***	0.176***	0.198***	0.141***
		(0.036)	(0.042)	(0.033)	(0.022)	(0.021)
C) Implied employment elasticity with respect to benefits (ε)						
		ε_{TypeA}	ε_{TypeB}	$p(\varepsilon_{TypeA} = \varepsilon_{TypeB})$	$\varepsilon_{TypeA} / \varepsilon_{TypeB}$	
Males	1	-1.062	-0.669	<0.001	1.587	
	2	-0.640	-0.684	0.197	0.936	
Females	1	-1.781	-0.528	<0.001	3.373	
	2	-1.372	-0.408	<0.001	3.363	
D) Implied labor earnings elasticity with respect to benefits (ε) at the household level						
		ε_{TypeA}	ε_{TypeB}	$p(\varepsilon_{TypeA} = \varepsilon_{TypeB})$	$\varepsilon_{TypeA} / \varepsilon_{TypeB}$	
Household leve	1	-1.259	-0.277	<0.001	4.545	
	2	-0.829	-0.374	<0.001	2.217	
Observations		4,072	4,072	4,072	4,072	

Note: The table shows the estimated effects of being granted residency after the reform on labor market outcomes by family type. Standard errors in parentheses. Column (5) is computed by creating a dummy $Emp = 1$ if one or both spouses are employed, and we then estimate Equation (2) while controlling for Emp and Emp interacted with the Type A and Type B dummies. Panel A shows, by gender, the effects of both spouses being granted residency after the reform relative to both spouses being granted residency before the reform on subsequent probability of being employed, unemployed, or not in the labor force, and/or dropped out of the welfare system completely for a balanced sample of couples in year 1 and 2 since LAS's residency. Panel B shows the same for couples where only one spouse is granted residency after the reform relative to both spouses being granted residency before the reform. Results in panels A and B are estimated as by Equation (2). Panel C shows implied employment elastiticies with respect to benefits levels estimated as the percentage change in employment relative to the percentage change in potential benefits levels induced by the reform. Panel D shows implied labor earnings elastiticies with respect to benefits levels estimated as the percentage change in labor earnings relative to the percentage change in potential benefits levels induced by the reform (see Table A.5 for full set of estimates for labor earnings and transfers).

* p<0.1; ** p<0.05; *** p<0.01

Table 6. Effect of reform on number of crime convictions 1 and 4 years following residency (accumulated from date of residency) for all adults and for couples.

	Year 1				Year 4			
	All adults	Couples	Males in couples	Females in couples	All adults	Couples	Males in couples	Females in couples
A) All crime								
Reform	0.029**	0.028**	0.031	0.026**	0.053***	0.067***	0.104***	0.045*
	(0.012)	(0.011)	(0.022)	(0.012)	(0.018)	(0.018)	(0.034)	(0.024)
Pre reform mean	0.024	0.012	0.022	0.012	0.096	0.037	0.102	0.056
Native Danes	0.013	0.005	0.010	0.002	0.053	0.022	0.045	0.010
B) Property crime								
Reform	0.031**	0.028**	0.028	0.027**	0.049**	0.056***	0.073*	0.046*
	(0.013)	(0.012)	(0.023)	(0.012)	(0.019)	(0.020)	(0.042)	(0.024)
Pre reform mean	0.019	0.011	0.016	0.012	0.075	0.034	0.076	0.056
Native Danes	0.007	0.002	0.004	0.002	0.024	0.009	0.016	0.007
C) Shoplifting from supermarkets								
Reform	0.022**	0.023***	0.018	0.026***	0.032**	0.035**	0.033	0.036*
	(0.009)	(0.008)	(0.012)	(0.009)	(0.014)	(0.016)	(0.028)	(0.020)
Pre reform mean	0.011	0.005	0.009	0.010	0.045	0.023	0.045	0.045
Native Danes	0.002	0.001	0.001	0.001	0.007	0.003	0.003	0.003

Note: Table shows the reform estimates on and pre reform means of number of crime convictions, by crime type and years since residency. Estimates are based on accumulated number of crime convictions since residency (and not number of crimes in a given year). Panel A shows results for all crimes, Panel B shows results for property crimes, and Panel C shows results for shoplifting from supermarkets. Results for 'All adults' are based on the base sample of adults (age 18-55 on the time of residency). Results for 'Couples', 'Males', and 'Females' are based on 18-45 year old individuals from the couples sample. The corresponding number of crime convictions from 2002, and 1 and 4 years ahead, respectively, for native Danes are obtained from sampling the full population of adults with the same age range, gender, and household type.
* p<0.1; ** p<0.05; *** p<0.01

Table 7. Effect of reform on children age 0-5 at residency, probability of enrollment in childcare or preschool, and Danish test scores during school-ages.

Outcome	Pre reform mean	Estimate	Observations
A) P(Childcare / prescool use)			
Year 1 since residency	0.558	-0.157**	950
	(0.497)	(0.058)	
Year 2 since residency	0.679	-0.144**	676
	(0.468)	(0.066)	
B) Test scores, Danish, first recorded test following residency			
Total test scores, standardized relative to national average	-0.676	0.056	875
	(0.953)	(0.122)	
P(2.5sd below national average or lower), Total test score	0.030	0.052**	875
	(0.171)	(0.022)	
------ , Language comprehension subset	0.010	0.035**	875
	(0.099)	(0.016)	
------ , Decoding subset	0.034	0.027	875
	(0.153)	(0.029)	
------ , Reading comprehension subset	0.030	0.020	875
	(0.171)	(0.021)	

Note: The table shows the estimated effects of being granted residency after the reform relative to the individual being granted residency before the reform for children between age 0 and 5 at time of residency. Panel A shows effects on subsequent probability of the child being enrolled in childcare/preschool in year 1 and 2 since residency (for year 2 we only consider those who are 0-4 at residency as those who are 5 at residency have started kindergarten). Panel B shows effects on total language test scores in the national test in Danish schools, and corresponding probabilities of falling at or below 2.5 standard deviations below the national average together with the three underlying subsets (language comprehension, decoding, and reading comprehension) that make up the aggregate test score. Test scores are measured at different grades-levels as children's time of residency and age at residency vary, and the national test scores are only available from 2008 and onwards. 19% are measured in 2nd grade, 41% in 4th grade, 33% in 6th grade, and 8% in 8th grade. Not all children have recorded test scores (due to absence at the day of test or special needs school / migrant class enrollment). Test for significant differences in missing test scores around the reform: P-value=0.223. Results are robust to limiting the sample to test scores measured in grades 2, 4, or grades 2, 4, and 6, or limiting the sample to individuals age 0-4 at time of residency.
* p<0.1; ** p<0.05; *** p<0.01

Table 8. Effect of reform on children's overall years of schooling.

	Pre reform mean	Mean effect	Males	Females
Reform, all	11.44	-0.093	-0.171	-0.005
	(1.75)	(0.103)	(0.180)	(0.169)
Reform, age at residency<10	11.10	0.083	0.157	0.039
	(1.41)	(0.138)	(0.190)	(0.126)
Reform, age at residency>=10	12.27	-0.647**	-0.897***	-0.270
	(2.17)	(0.267)	(0.292)	(0.422)

Note: The table shows estimated effects of being granted residency after the reform on years of completed schooling measured in 2016 for children who are between age 2 and 13 at timing of residency. The first column presents pre reform means. The second column presents mean results for the full group, and columns 3 and 4 present results for males and females, respectively. Estimates are evaluated relative to receiving recidency before the reform. The first row shows estimates for all children age 2-13. In the other rows, the estimated effect is presented across age at residency (younger than 10 years / 10 years or older). Of the 2,648 children age 2-13, 1,409 are males and 1,239 are females. * p<0.1; ** p<0.05; *** p<0.01

Table 9. Effect of reform on number of crime convictions at or before age 20 for children age 11-18 at residency.

Treatment	All crime	Property	Theft from supermarket	Other shoplifting	Violence
Pre reform mean	0.407	0.261	0.058	0.025	0.102
	(0.993)	(0.719)	(0.276)	(0.176)	(0.412)
A) Full sample	0.269*	0.123	0.001	0.036	0.078***
	(0.151)	(0.136)	(0.040)	(0.023)	(0.025)
B) By gender					
Males	0.416	0.181	0.027	0.016	0.140***
	(0.256)	(0.222)	(0.046)	(0.027)	(0.050)
Females	0.041	0.026	-0.016	0.056**	0.007
	(0.078)	(0.076)	(0.042)	(0.019)	(0.011)
C) By age at residency					
Age at residency < 14	0.155	0.014	-0.049	0.002	0.087
	(0.281)	(0.230)	(0.068)	(0.025)	(0.061)
Age at residency >= 14	0.444**	0.254	0.051	0.061*	0.114**
	(0.189)	(0.156)	(0.039)	(0.033)	(0.051)
Observations	1,025	1,025	1,025	1,025	1,025

Note: The table shows the estimated effects of being granted residency after the reform relative to being granted residency before the reform for children between age 11 and 18 at time of residency on subsequent number of crime convictions until age 20 for (Panel A) all children age 11-18 at time of residency, (Panel B) by gender, (Panel C) by age at residency (younger than 14, 14 or older). All estimates are evaluated relative to receiving residency before the reform.
* p<0.1; ** p<0.05; *** p<0.01