Do Vocational Language Courses Improve Refugee Integration? Evidence from Germany

David Titus*

Working Paper †

July 5, 2026
(View the Most Updated Version Here)

Abstract

I estimate the dynamic impact of the German Vocational Language Course program (Berufssprachkurse or BSK, and formerly known as the ESF-BAMF course) on the employment and language skills of refugees in Germany. I further investigate whether the returns to the course are impacted by anti-refugee prejudice, and I also test for spillover effects within refugee households. Finally, propensity scores are used to generate overlap weights and estimate effects on the overlap population of refugees with similar treatment propensities. In the end, my event study results imply that the BSK increases fluency in German by 0.10 to 0.15 standard deviations compared to refugees who took only the basic language course. Furthermore, I find that participation is associated with substantial increases in employment probability that may become stronger several years after taking the course. There is also suggestive evidence that increased prejudice may reduce the marginal return to BSK. While there are only limited indications of within-household spillovers to German fluency, I do find evidence that refugees become less likely to be employed if someone else in the household partakes in BSK. These results imply that vocational language courses have remained an effective tool in improving refugee language acquisition and employment despite the large and sudden increase in asylum seekers from 2015 to 2018.

*PhD Candidate, Department of Economics, Cornell University, dwt45@cornell.edu

†I would like to thank Francine Blau, Michèle Belot, Lawrence Kahn, Evan Riehl, Julia Lang, and all others who have provided helpful feedback and guidance. Any mistakes are my own.

I. Introduction

Language and vocational training courses are considered a vital policy tool for facilitating refugee integration. Many countries in Europe have devoted considerable resources to such courses in the wake of large refugee inflows. Such courses will likely continue to be of major importance given the significant increase in refugees who have arrived over the past decade from Syria, Ukraine, Myanmar, Afghanistan, Sudan, and other nations. For that reason, it is important to ensure that integration courses have remained effective as they have expanded due to the large and sudden influx of refugees from unexpected countries.

While many prior researchers have found introductory language courses in specific countries to be a reliable tool in improving refugee fluency, labor market outcomes, and economic welfare (Arendt et al. 2021; De Vroome and van Tubergen 2010; Kaida 2013; Lochmann et al.2019; Kanas and Kosyakova 2023), less research has focused on whether refugee language courses with vocational training components have remained effective. Part of the reason for this is the difficulty in identifying a true causal effect: unlike basic language courses that are ostensibly required for refugees with insufficient fluency, more advanced courses tend to be voluntary, and acceptance into the course may be at the discretion of a caseworker with relative leeway in decision-making (Lang 2022). Their take-up is thus often heavily affected by both unobservable refugee characteristics and local factors that may simultaneously increase take-up and improve refugee labor market prospects. No prior paper, to my knowledge, has examined whether more specialized vocational language courses produce spillovers to household members or whether these courses are less effective in regions with higher prejudice.

Germany is an ideal setting to analyze these questions: between 2015 and 2020, over 800 thousand Syrians sought refuge in Germany, with tens of thousands more arriving from other countries. Newly arrived refugees were initially assigned to states and districts according to administrative allocation formulas based largely on population and tax revenue. They were also initially prohibited from taking up residence in different states for up to a year. Some states have even imposed additional three-year residence restrictions at the district level. These initial residency restrictions mean that refugees are randomly exposed to exogenous variation in local integration resources and course availability and are also less apt to selectively migrate out of their initially assigned region.

In this paper, I estimate the effect of Germany's Vocational Language Course (Berufssprachkurse or BSK, and formerly organized as the equivalent ESF-BAMF course before 2017) on refugee fluency in German and the probability of employment. I also test whether having a family member take the course affects the outcomes of household members, and I further investigate whether prejudice affects the course's marginal returns. In order to account for potentially unobserved factors that may be correlated with course participation, I exploit panel data and use a fixed-effects model to estimate both overall effects and dynamic effects over time in a stacked event study. Finally, I use propensity scores to apply overlap weights to my previous methods. This approach simultaneously makes treatment and control groups more comparable and estimates the effect of BSK participation for those with similar treatment propensities. As I will later discuss, this is an ideal estimand for

policymakers to consider, as the BSK is a targeted intervention for certain refugees rather than a universal program.

I ultimately find that among refugees, this vocational language course increases German language skills. The effect size varies by empirical strategy, but it is about 0.06 standard deviations using a fixed-effects model and 0.10 to 0.15 standard deviations using a stacked event study. Employment only increases by 5 percentage points in my most basic specification, but the estimated effect size several years after course completion is as high as 29 percentage points when I use a stacked event study with overlap weights.

There are only marginal indications that the fluency in German of household members increases with an individual's participation. On the other hand, the employment of other household members seems to be negatively affected, implying a substitution of labor between the course participant and family members. Finally, there is suggestive evidence that higher levels of regional prejudice are associated lower BSK returns to employment, although this result is not precisely estimated.

This paper contributes to the literature in four ways. First, it estimates long-run dynamic returns to vocational language courses and provides evidence that their effects persist well beyond course completion. Second, it suggests that local prejudice may attenuate the benefits of integration programs. Third, it is the first paper to my knowledge to analyze potential refugee household spillovers from vocational courses. Fourth, it addresses crucial identification concerns by combining overlap weighting with individual fixed effects, thereby accounting for unobserved time-invariant factors while estimating a treatment effect for those with similar treatment propensities.

II. Background and Institutional Setting

Initial Refugee Settlement within Germany

In Germany, new asylum seekers are assigned to reception centers according to the Königstein Key (Königsteiner Schlüssel), a formula designed to allocate new refugees. Specifically, a given state's allocation share of any refugee cohort is equal to one third of its German population share multiplied by two thirds of its tax revenue share two years prior (Albarosa and Elsner 2023). Refugees are then further randomly sorted into smaller districts and municipalities, usually in proportion to each region's population share.[1] Before 2016, refugees would initially remain in centers for up to six months while their cases were processed before being permitted to take up residence elsewhere. However, since the passage of the Integration Act of 2016, refugees have only been allowed to take up residence in their state of initial assignment for their first three years in Germany. In addition, seven of the sixteen federal states in Germany passed more strict residency restrictions, limiting refugees to their initially

[1]The states do have discretion in choosing how to allocate new refugee arrivals among smaller districts, but they in practice do so according to district-level population most of the time (Albarosa and Elsner 2023).

assigned local region for some or all of that period.[2] The initial assignment process limits selective geographic sorting and helps ensure that refugees face plausibly exogenous variation in local integration resources and course availability.

Integration Courses and New Refugee Cohorts

While refugees are allowed to take up employment after three months in Germany, the Federal Office for Migration and Refugees (BAMF) provides a mandatory language training course, known as the BAMF Integration Course, for refugees with insufficient German fluency. This is a 600 hour course and it takes six to seven months to complete. Upon completion, refugees are given a language test and are given different certificates commensurate with their fluency levels. Even though this is supposed to be a mandatory course for all refugees who do not know German, exceptions can exist for health reasons or if the refugee has become employed and does not have the time to take the course. Because of this, significant geographic variation in course take-up can exist. In Figure 1, I use administrative data from the BAMF to generate a heat map of the number of new integration course takers in 2016 as a fraction of the number of refugees entitled to take the course by NUTS3 region (also referred to as district or Landkreise). As shown, there is significant variation by district, and there is no clear geographic pattern in take-up.[3]

In addition to the BAMF Integration Course, the BAMF also offers a supplementary course to job-seeking refugees, typically offered to those who have been certified through the BAMF Integration Course with at least a beginner's level of fluency in German. This program is currently referred to as the German for Vocational Language Course (Berufssprachkurse or BSK in German), formally called the ESF-BAMF Course.[4] The course lasts from six months (with a full-time daily commitment) to a year (with a part-time daily commitment) and is designed to teach technical language, resume building, and professional skills. Toward the end of the course, refugees are matched with employers to shadow other employees and hopefully make connections that will lead to them being hired. In addition to having to demonstrate sufficient language skills, only unemployed, job-seeking refugees may apply for this training in most cases. Their acceptance is at the discretion of local caseworkers who conduct interviews to determine eligibility. These caseworkers prioritize those most likely to benefit from the program but also are given some discretion in how to make that determination. The design of this procedure means that treatment is not randomly assigned and is likely correlated with both time-invariant and time-varying unobservables.

Before the large wave of Syrian refugees in 2015, the former ESF-BAMF course was associated with non-refugee students as often as refugee take-up. There has been some

[2]These states are Bavaria, Baden-Württemberg, North Rhine-Westphalia, Hesse, Saarland, Saxony and Saxony-Anhalt. See this 2018 report from the European Council of Refugees and Exiles for more details.

[3]Further variation also exists by the number of new integration courses run each year that cannot be fully explained by the number of refugees in the district: Manras and Kosyakova (2023) demonstrate that Eastern German regions run more integration courses, all else equal, than other parts of the country.

[4]Originally funded by the European Social Fund (ESF), the German government began funding the course themselves in 2016, resulting in the re-branding by the end of 2017. For the purposes of this paper, I will mainly refer to the course as BSK.

research on the BSK during the years before the large refugee inflow, and there have been indications of positive employment effects (Brücker et al. 2016, Lang 2022). However, causal identification has proven a significant hurdle. Lang (2022) took the most significant step forward in this endeavor: using administrative job center data in Germany, she studies the impact of BSK after tracking its 2014 participants over two years. To account for unobserved factors determining treatment take-up, she employs an instrumental variables approach. Namely, she exploits regional variation in the number of available courses and the number of unemployed immigrant job-seekers at each center to create a proxy for local treatment intensity. She ultimately finds that participants are nine percent more likely to be employed over a year after the course's completion. The downside is that this approach relies on regional variation that may also capture broader labor market conditions and local policy environments.[5] After all, both the number of unemployed foreigners will be correlated with local labor markets and local policy preferences. In any case, she does not analyze the new cohort of refugees.

Despite this earlier research, there is reason to suspect that external validity does not hold for the course's new participants. Beginning in 2015 and over just a couple years, more than 800 thousand Syrians, along with tens of thousands of refugees from countries in the Middle East and Central Asia, entered Germany. Not only do these refugees have very different backgrounds than economic immigrants, but their arrival led to changes in integration policy within Germany. In response to the large influx of refugees, the German government passed the Integration Act of 2016, which expanded the BSK, extended limits on migration within Germany to up to three years, and allowed states to further limit refugee residence to within districts and municipalities. Furthermore, refugees have faced high levels of prejudice in Germany, which may lower the course's effectiveness.[6] Jaschke et al. (2022) find that refugees living in more historically hostile regions culturally assimilate more quickly and are more likely to participate in voluntary integration and language courses. One interpretation is that refugees attempt to alleviate discrimination in more prejudiced areas by assimilating more quickly. An alternate interpretation is that local institutions in more prejudiced regions of Germany are more likely to enforce or encourage integration. Either way, prejudice has the potential to render BSK less effective for the newer refugee cohorts.

Baderschneider et al. (2024) are the first to attempt to identify the causal effects of BSK on these newer groups. By combining the IAB Establishment Panel with phone surveys, they use matching to compare BSK-treated and BSK-untreated refugees who completed the BAMF Integration Course in the same year. They ultimately find large positive effects for language and benefits for employment. However, they mainly evaluate effects for up to four years, so they do not observe dynamic employment effects several years after taking the course. A longer-run evaluation would be especially valuable, as it may take time for refugees to form networks as a result of the course.

[5]Lang (2022) demonstrates this by regressing her instrument on a large number of regional policy and market characteristics and finding significant correlations.

[6]In recent years, refugees have perceived increasing discrimination in both labor and housing markets in Germany. See Cumming and Heidinger (2025) for a more full discussion.

It is important to understand how these programs perform under the stress of a sudden refugee influx, whether their estimated effectiveness is long-lasting, and whether it is influenced by factors such as prejudice that may produce unique effects on refugees.

Refugee Outcomes: Language and Employment

Fluency in the language of the host country is an important skill for new immigrants to learn during their integration into the labor market. This is also true specifically for refugees: Chin and Cortes (2015) show that improvements in refugee wages are correlated with improvements in language. The magnitude of language's labor market impact is nearly impossible to consistently estimate due to unobserved ability and measurement error in self-reported fluency. That said, Dustmann and Van Soest (2002) find evidence that the downward attenuation bias from the latter far outweighs the upward bias from the former, implying that OLS estimates on the impact of language are significantly underestimated and that fluency is crucial for labor market outcomes.[7] Since language ability is likely a crucial mechanism in shaping immigrant labor market prospects, language-based integration courses are likely among the best policy tools to improve long-run outcomes for refugees.

Such tools are of particular importance in this setting. Even after adjusting for observable characteristics, a persistent residual refugee-native employment gap exists for refugees in Germany many years after immigration (Titus 2026). Part of the reason for this gap is likely a lack of vocational qualifications and training specific to Germany upon immigration. Since the BSK has a job search training and job shadowing program at the end, it may positively affect employment both directly through its vocational component and indirectly through its language component. That said, there are tradeoffs to these two parts: the longer a course focuses on language, the less it can focus on job training. Therefore, it is important to evaluate whether the BSK produces positive effects on language ability that meaningfully supplement the effects of the BAMF Integration course.

Household Spillovers

If the BSK course improves the language ability and labor market prospects of participating refugees, it may also affect the outcomes of other household members. Participants who acquire stronger German language skills may share those skills with family members, help them navigate German institutions, or provide information about employment opportunities and job search strategies. Improved labor market outcomes might also expand networks and increase access to resources that benefit other household members.

The direction of any employment spillover would be theoretically ambiguous. On the one hand, improved language ability and labor market information may facilitate employment among the partners of program participants. On the other hand, successful labor market

[7]It should be noted that not all studies that attempt to correct for bias find large effects. While it is nearly impossible to find a good instrument for language, Yao and Ours (2015) attempt to do so and conclude that fluency among immigrants in the Netherlands only impacts female wages and does not affect the probability of employment.

integration by the participant may, through an income effect, reduce the amount of work required by other household members. Either way, BSK courses have the potential to affect both the human capital of household members and intra-household labor supply decisions.

A relatively small body of prior work has documented household spillovers from vocational training programs. Most recently, Schlosser and Shanan (2025) find that an Active Labor Market Program (ALMP) in Israel led to an increase in employment and earnings among the partners of participants. There may also be human capital spillovers: Kugler et al. (2022) show that a vocational training program in Colombia improved the long-run higher education prospects of participants' children. Little is known about whether refugee integration courses can produce similar indirect effects, but these previous insights suggest that the BSK's benefits could extend beyond those of its participants.

III. Data

Sample of Interest

The 2015-2024 IAB-BAMF-SOEP Survey of Refugees, as part of the German Socio-Economic Panel, is a favorable dataset to analyze these questions. This is an over-sample of more recently arrived refugees in Germany that asks questions specific to issues concerning integration. In addition to refugee data on BAMF and BSK course participation and the year of immigration into Germany, the SOEP offers a panel component and links household-level identifiers with individual-level information on employment, age, education, marital status, number of children, and many other covariates. Also included are district-level measures of GDP, unemployment rates, and the foreign-born share of population. Employed individuals also have five-digit KLDB occupational codes, from which it is possible to determine the skill level of any employed observation. Finally, the survey contains data on self-reported ability to speak, write, and read German.

I am interested in the long-run effects of the BSK on the newer waves of refugees. I therefore restrict my sample of interest to refugees who were surveyed between 2015 and 2024 and who immigrated between 2015 and 2018.[8] Since the BSK is oriented toward the vocational training of working-age refugees, I limit my sample to only refugees ages 18 to 64.

Generating Variables

Fluency in German is one of my key outcomes of interest, so I must generate a measure of German language ability for refugees. I do this using three survey questions designed to capture self-reported fluency. These questions ask for a respondent's ability to read, write, and speak German on a scale of one to five. I standardize each of these responses into a

[8] I limit my sample to those who immigrated in 2018 or earlier so that any dynamic effects I estimate will not be affected by varying patterns among different cohorts. 2018 is a particularly desirable cutoff because very few surveyed refugees in the data arrived in 2019 or 2020. Most refugees in my final sample immigrated in 2015 and 2016.

standard normal distribution across the entire SOEP sample. I then take the average value of the standardized response across the three questions to generate my fluency measure.[9]

Since I am also interested in the potential effect of prejudice on the BSK's effectiveness, I generate a regional proxy for prejudice across the 401 German districts (the NUTS3 regions). I choose this level of aggregation since refugees are typically matched to BAMF Integration Course centers in their initially assigned district. To create a prejudice proxy that is unaffected by recent refugee cohorts, I follow Titus (2026) and use 2007 to 2014 SOEP data on native attitudes toward immigration. During these survey years, native respondents were asked on a scale from one to three whether they were worried about (1) immigration into Germany, and (2) hostility toward foreigners. After flipping the latter values so that greater concern of hostility positively correspond to greater worries about immigration, I standardize the survey responses. For each individual i and question k, I standardize each answer to a given question θ_{ikt} in year t, using 2007 survey responses as the benchmark mean and standard deviation:

$$\dot{\theta}_{ikt} = (\theta_{ikt} - \bar{\theta}_{k07})/\sqrt{var(\theta_{k07})}$$

After averaging the standardized response to each question, I take residual values from a regression of these responses on survey year fixed effects to account for secular trends in the responses over time. Finally, for each region over the pooled 2007-2014 period, I take the median level of prejudice within each NUTS3 region. Titus (2026) demonstrates that this proxy for prejudice is strongly correlated with an analogous measure that uses 2016-2023 survey questions to accurately capture 2016-2023 attitudes toward refugees.[10]

IV. Empirical Strategy

Fixed Effects Model

The most significant challenge in estimating the effect of vocational training courses is correcting for selection bias. Those who choose to partake in a vocational course often have very different observed and unobserved characteristics, meaning that a basic cross-sectional OLS regression would lead to biased results. This would certainly be the case here: as I will discuss in the Descriptive Statistics section, treated refugees likely have very different unobserved characteristics than untreated refugees. I therefore begin my analysis by accounting for time-invariant individual characteristics. Specifically, I exploit the panel structure of the SOEP and implement a fixed-effects estimation:

[9] If an observation is missing one or two of the three language measures, I record their fluency level according to the measures I have. For example, if a respondent reports written and spoken German ability but not reading ability, I take the average of the former two as my fluency measure.

[10] I have also tried using the median of the more contemporary prejudice measure, but the results are largely unchanged. Since the latter measure may be affected by refugee outcomes within each region, I elect to only use the lagged measure in this paper.

$$Y_{idt} = \alpha_i + \gamma_t + \lambda_1 BSK_{it} + \lambda_2 Prej50_{dt} \times BSK_{it} + \lambda_3 Prej50_{dt} + \mathbf{X'_{it}}\delta + \mathbf{Z'_{dt}}\rho + \varepsilon_{idt} \quad (1)$$

Y_{idt} is the outcome of interest (either fluency or employment) for an individual i in district d. BSK_{idt} denotes whether person i had already completed the course in year t. $Prej50_{dt}$ is the median level prejudice measure previously described in the Data section. This variable is time-varying since refugees can move between regions.[11] Note that the interaction effect between this measure and BSK completion reflects changes in BSK completion levels over time across regions, even for those who do not move. λ_2 may thus indicate on whether prejudice affects the returns to BSK. $\mathbf{X_{it}}$ is a vector of time-varying characteristics that includes a quadratic in age and years since migration, binary variables for secondary and post-secondary education, marital status, and number of children in the household. $\mathbf{Z_{dt}}$ includes district-level controls for GDP per capita, unemployment rates, and foreign-born population shares.[12] Importantly, I include individual fixed effects α_i, which will absorb any time-invariant individual characteristics that may be correlated with course participation.

My two coefficients of interest in this specification are λ_1 and λ_2. The most important of these two for this study is λ_1, which represents the overall post-treatment effect of BSK participation. λ_2, meanwhile, captures whether the returns to the course differ systematically between more and less prejudiced regions.

Since the BSK is primarily targeted to refugees who have already taken the introductory BAMF Integration Course and passed with sufficient proven language skills, it may be prudent to further limit my sample to only refugees who took the BAMF course before taking the BSK. In addition, it may be interesting to observe the effect of BSK itself. I therefore run alternative specifications. The first is run on the full sample but includes a binary indicator of whether someone has completed the BAMF course. The second is only run on those who completed the introductory course prior to treatment.

There is also a concern that my estimated effects could be biased by selective out-migration. If potential participants can predict which districts will be most accommodating in allowing them to take the BSK, they may selectively sort into those districts. I address this concern by running a different set of specifications in which my sample is limited to the seven federal states that restrict refugee residence to the district level during their first three years since migration.

[11] It should be cautioned that the estimated effect of prejudice on its own in this regression is not particularly informative: since I control for individual fixed effects, prejudice is identified by moving refugees. Thus, this estimated effect may be partially driven by selective migration.

[12] I could alternatively include district fixed effects, but I am interested in the effect of prejudice, which is time-invariant across districts. Furthermore, I already control for individual fixed effects. Titus (2026) shows that refugee migration between regions is relatively uncommon between consecutive years. In any case, alternative specifications in which I include district FEs yield similar results (available upon request).

Stacked Event Study

It is important in this setting to estimate not only the overall effectiveness of the BSK, but also dynamic effects. This is especially vital for evaluating its effect on employment: since course participants are supposed to not be employed during participation, their transition into the labor market upon completion may take time, and there may even be an initial negative effect immediately following course completion.

To evaluate the dynamic effects of the BSK while avoiding issues with staggered treatment timing, I implement a stacked event study (Cengiz et al. 2019). I do so using an event window ranging from 4 years before to six years after starting the course:[13]

$$Y_{idst} = \alpha_{is} + \gamma_{st} + \sum_{\substack{k=-4 \\ k\neq -1}}^{6} \beta_k \mathbf{1}\{t - g_i = k\} + \mathbf{X'_{ist}}\delta + \mathbf{Z'_{sdt}}\rho + \varepsilon_{idst} \tag{2}$$

Here, g_i indicates the treatment year of a given refugee i. Since I only have data on the year of BSK completion and the course takes up to a year, I define g to be the year of course completion minus 1 so that the event year is the year the refugee likely began the course.[14] For each treatment year g, I create a "stack," a cohort-specific sample consisting of refugee observations who are (1) untreated across the stacked event window, (2) not-yet-treated and observed before year g, and (3) complete the course in year $g+1$. I then pool all these stacks together for the above estimation. s denotes the stack to which an observation belongs. This approach ensures that treated observations are never used as controls for any cohort. α_{is} and γ_{st} are stacked individual and year fixed effects, respectively.

The key identifying assumption of this model is that untreated and treated observations would have followed parallel trends absent treatment. This is a rather strong assumption in this setting: it is very possible that treated refugees choose to partake in BSK in response to deteriorating labor market prospects or unobserved shocks. If so, treated refugees may exhibit a downward-sloping trend in outcomes prior to treatment, implying that the estimated post-treatment effects would be biased downward. To assess the parallel-trends assumption, I can examine the pre-treatment event-time coefficients and observe whether treated and control groups followed similar trends before the course was taken.

Household Spillovers

To evaluate whether BSK participation affects the language ability and employment of adult household members, I implement the prior two empirical approaches, except I drop any

[13] I do not include a larger range of event years due to a lack of sufficient observations. These early and late ever-treated observations are not included in the regression, nor in my descriptive statistics tables.

[14] I choose this as the event year because course participation is mechanically associated with unemployment. This means that if I choose the year of completion as the event year, the estimated employment event time coefficients would all be biased upward since there will inevitably be a strong negative treatment effect in omitted period $t = -1$. Similarly, the fluency effects would be biased downward since refugees should begin improving in language ability when they begin the course.

post-treated observation, and I now define treatment as whether a member of household h has completed BSK. Accordingly, equations 1 and 2 become:

$$Y_{ihdt} = \alpha_i + \gamma_t + \lambda_1 HHBSK_{it} + \lambda_2 Prej50_{dt} \times HHBSK_{it} + \lambda_3 Prej50_{dt} + \mathbf{X}'_{it}\delta + \mathbf{Z}'_{dt}\rho + \varepsilon_{ihdt} \quad (3)$$

$$Y_{ihdst} = \alpha_{is} + \gamma_{st} + \sum_{\substack{k=-4 \\ k\neq -1}}^{7} \beta_k \mathbf{1}\{t - g_h = k\} + \mathbf{X}'_{ihst}\delta + \mathbf{Z}'_{sdt}\rho + \varepsilon_{ihdst} \quad (4)$$

$HHBSK$ is equal to one if a household member has completed the BSK, and g_h is the year before that person completes the course. Note that for household spillovers, I have sufficient data to include a seventh post-treatment year.

Overlap Weighting

Neither of the prior two strategies fully addresses concerns that treated and untreated refugees are not comparable. After all, BSK courses are specifically designed for a certain kind of refugee who stands to benefit the most from the course. Therefore, to improve comparability between treated and untreated observations, I implement an overlap weighting approach (Li et al. 2018) by generating propensity scores from pre-treatment characteristics and pre-treatment outcomes. Namely, I first include only one unique observation per refugee and estimate propensity scores using a logit model:

$$\hat{p}_i = Pr(EverBSK_i = 1 \mid \mathbf{X_{bi}}) = \frac{\exp(\theta_0 + \mathbf{X_{bi}}'\delta + \mathbf{Z_{bi}}'\rho)}{1 + \exp(\theta_0 + \mathbf{X_{bi}}'\delta + \mathbf{Z_{bi}}'\rho))}, \quad (5)$$

θ_0 is a constant term, and b denotes baseline, pre-treatment characteristics in the first year surveyed. I define these as the age and years since migration in the first year a refugee was surveyed, along with the pre-treatment mean employment and language fluency of refugee i. Also included are all other pre-treatment baseline characteristics in equation (2), along with a binary indicator for gender. For never-treated refugees, the mean fluency and employment variables are equal to their mean characteristics across all their individual observations in the sample.

To generate unbiased propensity scores, the key assumption is that treatment assignment is independent of potential outcomes conditional on observables. This assumption holds in this setting as long as the caseworker's decision on whether to enroll a refugee is exclusively correlated with observable covariates that influence outcomes, such as language ability, employment status, and prior education. Naturally, this assumption may not hold given caseworkers' specific knowledge of each applicant, but since factors like prior BAMF integration course participation and employment status are key factors in determining eligibility, my observable covariates may provide reasonable treatment predictions.

From this estimation, one can generate a propensity score $\hat{p}_i$ for each refugee conditional on pre-treatment characteristics. This score ideally should predict the probability that a given refugee is ever treated at any point across the 2015-2024 sample window.

To create overlap weights, one can simply use the propensity score itself for never-treated observations. For ever-treated observations, I subtract p_i from 1. The idea is to down-weight any untreated observation who is either very unlikely to be treated, or any treated observation who is very likely to be treated. I then repeat my prior two estimation strategies using these weights. The resulting estimation yields a treatment effect on an overlap population of treated and untreated observations with similar treatment propensities.[15]

It should be stressed that the estimand of interest in this approach is not the average treatment effect (ATE) on the entire refugee population, but rather the average treatment effect on the overlap population (ATO). In other words, overlap weighting emphasizes treated refugees who resemble non-participants and non-participants who resemble participants. The resulting estimates reflect the effect of BSK for refugees who have similar treatment propensities based on observable characteristics. Such an estimand is particularly relevant in this context: the BSK is not designed to help all refugees, but rather to improve labor market integration among otherwise skilled refugees who face employment barriers. Since the overlap population reflects refugees with treatment propensities similar to those observed among both participants and non-participants, this estimated effect may be particularly informative for policymakers considering whether to expand BSK enrollment.

V. Descriptive Statistics

I begin my descriptive analysis by providing an overview of the years in which refugees entered Germany and the years in which the BAMF Integration Course and the BSK were completed. These statistics are displayed in Table 1. Notice first that only 3,412 out of 19,829 observations (about one in six) in my sample have taken the BSK, indicating that it is reserved to only a particular set of refugees. As shown, most refugees in my sample immigrated in 2015, and it typically takes refugees one to two years to complete the introductory BAMF course and two to three years after migration to complete the BSK. In addition, my sample contains a diverse distribution of observations by years since migration. This should be expected, given that refugees are interviewed in follow-ups following their initial entry into the survey.

I next analyze the differences in observable characteristics between treated and untreated refugees. To do so, I provide summary statistics of important variables by treatment status in Table 2. As shown, never-treated observations have similar language skills to pre-treated refugees, but they have far lower fluency than post-treated refugees. They are also much more likely to be employed than pre-treated refugees: before taking the course, only 9.3 percent of pre-treated refugees are employed, less than a third of the never-treated mean. This is to be somewhat expected given that one must be unemployed to qualify for the

[15] In addition to the relevance of this particular estimand, this approach also does not have to make arbitrary trimming decisions, as would be required when using inverse probability weights.

course. In addition, both the pre-treated and post-treated refugee groups have about ten percentage points of higher average education than non-treated refugees. Again, this reflects the fact that the BSK is oriented to those with high ability that would benefit the most from vocational training. As expected given that it is in most cases a prerequisite, treated observations are more likely to have completed the BAMF integration course: among post-treated observations, 90 percent have done so, whereas only 58.4 percent of never-treated observations have done the same.

Although there are too few observations for meaningful estimation, I also include summary statistics for real hourly wages and for whether a refugee is in a high-skill occupation.[16] Interestingly, refugees who are never treated seem to have higher hourly wages than both pre-treated and post-treated refugees, although the limited sample size makes the estimate imprecise. This may reflect the fact that those who did not take the course may have found employment sooner and have thus accumulated greater on-the-job human capital.

There does not seem to be an obvious correlation between the lagged median prejudice measure and the probability of taking the course. While this does provide evidence for exogenous variation for my prejudice effect estimates, it also indicates that integration resources such as BSK are equally accessible across prejudiced and unprejudiced regions.

Table 3 reports the same summary statistics as Table 2, but it restricts the sample to only refugees who have taken the BAMF Integration Course. This table therefore offers comparative statistics between refugees who are likely much more comparable in their eligibility to be enrolled in BSK. Even so, the patterns in the descriptive statistics summarized above largely persist. Similar patterns are also apparent if I further restrict the sample to refugees living in the seven states with more strict residency restrictions.[17]

Overall, BSK participants have noticeably different pre-treatment characteristics from non-participants. Even though I can control for many of these differences, including unobserved time-invariant characteristics, it is very possible that they will also differ in time-varying factors that both cannot be directly observed and can affect treatment adoption.

VI. Results

I next analyze the results of my various empirical specifications. In total, I find a strong positive effect of BSK on German fluency, as well as strong indications of a positive effect on employment.

[16] Hourly wages are calculated by dividing monthly earnings by actual hours worked per week and by 4.3. Values are winsorized at one third of the first percentile and three times the 99th percentile and are adjusted for inflation to 2012 Euro. Skill level is binary and equal to one for values with a KLDB occupational code with a final digit of 2, 3, or 4.

[17] These descriptive statistics can be found in Appendix Tables A1 and A2.

Fixed Effects Model

I report the results from the fixed effects model described in equation (1) in Table 4. To sum up, these estimates indicate that completion of the BSK is positively associated with a 0.068 standard deviation increase in fluency (see column 1), along with a 7.5 percentage point increase in employment probability among the full refugee sample (see column 3). Among refugees who have completed the BAMF Integration course, the fluency effect is virtually identical at 0.065 (column 5), but the employment effect drops to 0.031, with a p-value above 0.15 (column 7).

It appears from the table that regions with higher prejudice may have slightly lower returns to employment. To interpret these relevant coefficients, keep in mind that the standard deviation of the prejudice measure is about 0.11. Thus, a one standard deviation higher level of median prejudice is associated with about a 1.57 percentage point decrease in the BSK return to employment for the full sample, and a 1.99 percentage point lower return to employment among all BAMF completers (columns 4 and 8). That said, these results are not quite statistically significant, although they are similar in magnitude and sign across both specifications. Prejudice's effect on the course return should only be taken as suggestive.

The estimated effects on fluency and employment actually appear stronger when I restrict the sample to the seven more restrictive residency states. The results from estimating equation (1) on this more limited sample are shown in Table 5. The results from this table reveal that BSK completion is associated with a 0.120 SD increase in fluency (column 1) and a 8.6 percentage point increase in employment (column 3) among the full sample. In the BAMF completers sample, these estimated effects are 0.094 and 3.6 percentage points, respectively (columns 5 and 7). There is again a negative association between prejudice and course effectiveness, although it is again not statistically significant at the 10 percent level.

In total, these initial results imply that the BSK has produced sizable positive overall effects for the 2015-2018 refugee cohort.

Stacked Event Study and Dynamic Effects

I next turn to my stacked event study from equation (2) to evaluate dynamic effects over time. In so doing, I can also test whether treatment and control observations, conditional on observable covariates and time-invariant characteristics, follow similar pre-trends. If so, then that would lend credence to my overall estimates from the previous section.[18]

I display the event study results for fluency and employment in Figure 2.[19]. For fluency, there is no pre-trend, and the estimated post-treatment effects are even larger than my fixed-effects model would have implied. Namely, there appears to be roughly a 0.10 to 0.15 SD increase in fluency relative to the control group for the first three years after course completion. This pattern holds for both the full sample and the BAMF completer sample,

[18]Recall, however, that for these results, the event year is defined as the year a refugee starts the course, proxied by the completion year minus 1

[19]An analogous figure for the more restricted set of states can be found in the Appendix (see Figure A1)

indicating that the BSK produces additional benefits to language ability beyond the basic course.[20]

However, Figure 2 also reveals a persistent downward-sloping pre-trend for employment. Specifically, pre-treated course takers appear to have increasingly worse employment prospects relative to untreated individuals as they near the treatment year. This pre-trend does not initially appear to reflect a temporary shock, as it is consistent throughout the pre-treatment event window. Instead, this is likely caused by systematic differences between refugees who choose to wait several years after migration to take the BSK course, and those who do so relatively soon after taking the basic course: refugees who choose to not take the course soon likely do so due to better early labor market outcomes. In any case, this downward pre-trend completely reverses in post-treatment years: four years after course completion, the estimated treatment effect on employment rises to between 11 and 18 percentage points, depending on the event year and sample choice. A plausible story could be that otherwise highly skilled refugees only apply to take the BSK if they face worsening labor market outcomes.

Household Spillovers

Since the BSK appears to significantly increase employment and fluency, I next investigate whether the program produces spillovers to adult household members who have not taken the course. I begin by implementing the fixed-effects model in equation (3) on the 18-64 sample of non-BSK-takers. The results of this implementation can be found in Table 6.

These estimations provide initial indications that having a household member take the course leads to greater fluency but lower employment for the other adult household members. Specifically, such a treatment is associated with a 0.099 SD increase in fluency but an 11.0 percentage point decrease in employment. While it might initially seem unclear why the course would lead to negative employment spillovers, it is very possible that this is due to an income effect. Suppose, for instance, that the BSK course-taker becomes the breadwinner in the household due to the program. In that case, other household members may have less incentive to work due to the increased income from the course-taker.

That said, the stacked event study strategy casts doubt that there is a true language spillover. Figure 3 displays the dynamic effects for non-BSK household members, and there is a positive pre-trend for fluency among household members. It is thus unclear whether the program is actually producing a spillover in language ability. There is, however, no pre-trend for employment, implying that there may be a true causal impact. The effect for employment also persists even seven years after course uptake.

In summary, my results indicate that there is only limited evidence for positive fluency spillovers, while there are stronger indications of negative employment spillovers.

[20] The full numerical results for all event-time coefficients can be found in Appendix Table A3.

Overlap Weighting

My prior results have given strong indications that language ability is positively affected by BSK. However, my stacked event study reveals that the employment results, while statistically significant and sizable, may be biased downward. Part of the issue is likely that my treatment and control groups are not comparable enough. Ideally, I would want to compare a treated refugee on the margin of being treated with an untreated refugee also likely to be treated. This is why I conduct overlap weighting. In this section, I re-run my previous estimations on course-takers after generating overlap weights for all observations.

I first report the marginal effect results of the logit model in equation (5) that predicted the propensity scores. The results of this estimation are in Table 7. From this table, we see, as expected, that pre-treatment employment and BAMF course completion are extremely strong predictors of course take-up: BAMF course completion at the baseline survey year is associated with a 4.9 percent increase in treatment probability. The average baseline employment coefficient is even larger at -0.188, indicating that average baseline employment is by far the strongest indicator of BSK participation. Gender and post-secondary education are also important predictors: women are 4.6 percent less likely to enroll in BSK, and those with post-secondary education are 4.9 percent more likely to enroll.

After adding the overlap weights, I redo my primary empirical strategies from equations (1) and (2). The results of the fixed-effects model from equation (1) can be found in Table 8. Overall, the fluency effects are not largely changed, except that the estimated effects on employment and fluency have consistently increased by 0.02 SD in the specifications with all observations. The employment results, meanwhile, are 2 percentage points larger across the board compared to the original Table 4 results and are statistically significant. Prejudice still is associated with lower course returns to employment, although the effect is still imprecisely estimated.

Overlap weights make the biggest difference for employment in the stacked event study. As shown in Figure 4, while there is still a pre-trend in employment, there appears to be a clear positive and growing treatment effect even six years after course completion. By year six, course-takers are 34.1 percentage points more likely to be employed than non-takers in the full sample and 31.6 percentage points more likely in the BAMF completer sample.[21] Course taker fluency also improves by 0.1 to 0.2 SD depending on the event year and specification.

Recall that these are not ATE estimates, but rather estimated average treatment effects on the overlap population. If these are true causal effects, and if class capacity for this course expanded by one seat, then the marginal new entrant, rather than a randomly chosen refugee, would see treatment effects of this magnitude.

[21] See Appendix Table A5.

VII. Conclusion

In this paper, I have estimated the effect of the German Vocational Language Course on refugee language ability and employment. I have done so using a fixed-effects regression, a stacked event study, and with overlap weights. I have additionally evaluated whether there are spillovers to other household members and whether increases in regional prejudice are associated with lower course returns.

I ultimately find evidence for a number of effects. First, the BSK course is highly effective at improving language ability for refugees, even among those who have already taken the basic language course. Second, the course appears to be strongly associated with improved employment outcomes for refugees even many years after completion. This is the first paper, to my knowledge, to find evidence of such lasting effects. Third, I have found suggestive indications that prejudice slightly lowers the employment returns to the course. This result suggests that the benefits of vocational courses may be larger in less prejudiced regions. Fourth, I have found that a household member taking BSK is associated with lower post-treatment employment among other household members, implying a positive income effect.

In addition, my overlap weighting approach yields a tractable interpretation for policy-makers who might consider a marginal extension in the program's capacity: refugees on the margin of enrolling into the program face potentially large benefits from such an expansion. Overall, these results imply that vocational language courses like the BSK play an important role in facilitating both human capital accumulation and labor market integration.

Tables and Figures

Table 1: Distribution of Refugees by Year of Immigration, Course Completion Year, and Years Since Migration

Year	Imm. Yr. (%)	BAMF (%)	BSK (%)	YSM	Percent (%)
2015	70.1	0.1	0.2	0	1.3
2016	19.0	21.8	7.6	1	15.0
2017	5.8	36.7	27.6	2	17.0
2018	5.1	17.7	33.1	3	14.2
2019	–	10.4	10.1	4	12.7
2020	–	8.3	10.6	5	13.8
2021	–	2.5	5.0	6	10.4
2022	–	1.3	–	7	7.1
2023	–	0.9	5.9	8	4.8
2024	–	0.2	–	9	3.7
Obs.	19,829	15,075	3,412	Obs.	19,829

Notes: Imm. Yr. is the year of immigration into Germany. BAMF and BSK indicate the years of BAMF Integration Course and BSK completion, respectively. YSM denotes years since migration.

Table 2: Descriptive Statistics by BSK Participation, Full Sample

	Never Treated		Pre-BSK		Post-BSK	
Variable	Mean	Obs.	Mean	Obs.	Mean	Obs.
Fluency	-0.599	15,050	-0.661	889	-0.120	2,499
Employed	0.313	16,426	0.093	889	0.443	2,514
Hourly Wage	13.255	3,831	10.010	42	11.853	903
High-Skill Occupation	0.639	4,325	0.597	74	0.693	925
Female	0.391	16,426	0.355	889	0.274	2,514
BAMF Course Completion	0.584	16,426	0.717	889	0.900	2,514
Secondary Education	0.095	16,426	0.054	889	0.091	2,514
Post-Secondary Education	0.181	16,426	0.280	889	0.324	2,514
Years Since Migration	3.945	16,426	2.208	889	4.634	2,514
Married	0.584	16,426	0.629	889	0.592	2,514
# of Children in Household	1.681	16,426	1.865	889	1.566	2,514
Lagged Median Prejudice	0.014	16,426	0.008	889	0.005	2,514

Notes: This table includes only observations surveyed from 2015 to 2024 and immigrated from 2015 to 2018. Language fluency is standardized from questions on German reading, writing, and speaking ability. Both education variables, Married, and Female are binary. BAMF Course is also binary and indicates whether the BAMF integration course has been taken. Hourly wage is in 2012 Euro. High-Skill Occupation is equal to one if the respondent has a KLDB skill level of at least 2, indicating at least some vocational or post-secondary qualification. Lagged median prejudice calculated from the weighted distribution on immigration attitude questions nine years before the survey year. There are 6,968 unique survey respondents.

Table 3: Descriptive Statistics by BSK Participation
Post-BAMF Course Only

	Never Treated		Pre-BSK		Post-BSK	
Variable	Mean	Obs.	Mean	Obs.	Mean	Obs.
Fluency	-0.463	9,512	-0.493	637	-0.067	1,946
Employed	0.341	9,596	0.119	637	0.468	1,957
Hourly Wage	12.604	2,482	10.010	42	12.026	781
High-Skill Occupation	0.485	2,732	0.501	67	0.688	762
Female	0.339	9,596	0.342	637	0.262	1,957
Secondary Education	0.083	9,596	0.061	637	0.083	1,957
Post-Secondary Education	0.198	9,596	0.295	637	0.360	1,957
Years Since Migration	4.182	9,596	2.568	637	4.812	1,957
Married	0.609	9,596	0.636	637	0.612	1,957
# of Children in Household	1.667	9,596	1.812	637	1.604	1,957
Lagged Median Prejudice	0.015	9,596	0.007	637	0.007	1,957

Notes: This table includes only observations surveyed from 2015 to 2024 and immigrated from 2015 to 2018 who have already taken the BAMF integration course. Language fluency is standardized from questions on German reading, writing, and speaking ability. Both education variables, Married, and Female are binary. Hourly wage is in 2012 Euro. High-Skill Occupation is equal to one if the respondent has a KLDB skill level of at least 2, indicating at least some vocational or post-secondary qualification. Lagged median prejudice calculated from the weighted distribution on immigration attitude questions nine years before the survey year. There are 3,791 unique survey respondents.

Table 4: Effects of BSK on Fluency and Employment

	All Refugees				BAMF Completers			
	Fluency		Employment		Fluency		Employment	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post-BSK	0.068**	0.066**	0.075***	0.076***	0.065**	0.063**	0.031	0.033
	(0.027)	(0.027)	(0.019)	(0.019)	(0.031)	(0.032)	(0.023)	(0.023)
Prej50 × BSK		0.207		-0.157		0.114		-0.199
		(0.197)		(0.140)		(0.235)		(0.154)
BAMF Course	0.206***	0.206***	-0.057***	-0.058***				
	(0.021)	(0.021)	(0.012)	(0.012)				
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	14,431	14,431	14,569	14,569	9,666	9,666	9,677	9,677

Notes: The dependent variables are standardized German fluency and employment status. Columns (1) through (4) use the full sample of refugees surveyed between 2015 and 2024 who immigrated in 2018 or earlier, while Columns (5)–(8) restrict the sample to refugees who completed the basic BAMF integration course. Prej50 refers to the lagged regional median prejudice level. All specifications include individual and survey-year fixed effects. Control variables include age squared, years since migration squared, educational attainment, marital status, number of children in the household, NUTS3 level unemployment rates, NUTS3 level GDP per capita, and the regional NUTS3 share of foreigners. Standard errors are clustered at the individual level and are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Effects of BSK on Fluency and Employment
Restricted Residency States Only

	All Refugees				BAMF Completers			
	Fluency		Employment		Fluency		Employment	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post-BSK	0.120***	0.111***	0.086***	0.095***	0.094**	0.125***	0.036	0.051*
	(0.035)	(0.038)	(0.024)	(0.026)	(0.041)	(0.044)	(0.028)	(0.031)
Prej50 × BSK		0.335		-0.398		-1.297*		-0.678
		(0.611)		(0.430)		(0.749)		(0.471)
BAMF Course	0.198***	0.198***	-0.069***	-0.069***				
	(0.025)	(0.025)	(0.015)	(0.015)				
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	9,393	9,393	9,488	9,488	6,254	6,254	6,264	6,264

Notes: The dependent variables are standardized German fluency and employment status. Columns (1) through (4) use the full sample of refugees surveyed between 2015 and 2024 who immigrated in 2018 or earlier, while Columns (5)–(8) restrict the sample to refugees who completed the basic BAMF integration course. Prej50 refers to the lagged regional median prejudice level. All specifications include individual and survey-year fixed effects. Control variables include age squared, years since migration squared, educational attainment, marital status, number of children in the household, NUTS3 level unemployment rates, NUTS3 level GDP per capita, and the regional NUTS3 share of foreigners. Standard errors are clustered at the individual level and are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: Household Spillover Effects

	Fluency		Employment	
	(1)	(2)	(3)	(4)
HH BSK	0.099*	0.092	-0.110***	-0.107***
	(0.057)	(0.057)	(0.034)	(0.035)
HH BSK × Prej50		-0.096		0.021
		(0.157)		(0.106)
Individual FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Obs.	11,399	11,391	11,534	11,526

Notes: The dependent variables are standardized German fluency and employment status. HH BSK refers to whether a family member in the household had taken the BSK course. Prej50 refers to the lagged regional median prejudice level. Only observations who have not taken the vocational training course are included. All specifications include individual and survey-year fixed effects. Control variables include a binary indicator for BAMF Integration Course completion, age squared, years since migration squared, educational attainment, marital status, number of children in the household, NUTS3 level unemployment rates, NUTS3 level GDP per capita, and the regional NUTS3 share of foreigners. Standard errors are clustered at the household level and are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 7: Propensity Score Logit for BSK Participation
Marginal Effects

	BSK Ever
Pre-treatment Fluency	0.020***
	(0.006)
Pre-treatment Employment	-0.188***
	(0.019)
Baseline Age	0.012***
	(0.003)
Baseline Age Squared	-0.000***
	(0.000)
Female	-0.046***
	(0.008)
Secondary Education	-0.019
	(0.018)
Post-Secondary Education	0.049***
	(0.008)
Baseline YSM	-0.020
	(0.013)
Baseline YSM Squared	-0.004
	(0.003)
Married	0.000
	(0.010)
# of Children in Household	0.002
	(0.003)
BAMF Course Completion	0.049***
	(0.008)
NUTS3 Unemployment Rate	0.002
	(0.049)
NUTS3 GDP per Capita	-0.000
	(0.002)
NUTS3 Foreigner Share	-0.001
	(0.001)
Obs.	5,488
Pseudo R^2	0.113

Notes: This table reports logit estimates of the probability of ever taking a BSK course. For ever-treated observations, pre-treatment fluency and employment are averaged over observations prior to BSK participation. For non-participants, these outcomes are averaged over all observed years. Only one observation per unique person identifier is included in the estimation. Baseline covariates such as age are measured in the respondent's first observed survey year. Standard errors are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 8: Effects of BSK on Fluency and Employment
Overlap Weighting

	All Refugees				BAMF Completers			
	Fluency		Employment		Fluency		Employment	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post-BSK	0.084***	0.087***	0.102***	0.092***	0.062*	0.056	0.060**	0.052*
	(0.029)	(0.031)	(0.021)	(0.022)	(0.034)	(0.038)	(0.026)	(0.028)
Prej50 × BSK		0.029		-0.102		-0.057		-0.067
		(0.089)		(0.068)		(0.114)		(0.082)
BAMF Course	0.180***	0.179***	-0.031	-0.026				
	(0.034)	(0.035)	(0.020)	(0.020)				
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	13,402	13,388	13,502	13,488	9,013	9,000	9,024	9,011

Notes: The dependent variables are standardized German fluency and employment status. Columns (1) through (4) use the full sample of refugees surveyed between 2015 and 2024 who immigrated in 2018 or earlier, while Columns (5)–(8) restrict the sample to refugees who completed the basic BAMF integration course. Prej50 refers to the lagged regional median prejudice level. All specifications include individual and survey-year fixed effects. Control variables include age squared, years since migration squared, educational attainment, marital status, number of children in the household, NUTS3 level unemployment rates, NUTS3 level GDP per capita, and the regional NUTS3 share of foreigners. Standard errors are clustered at the individual level and are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure 1: BAMF Integration Course Take-Up Rates, 2016

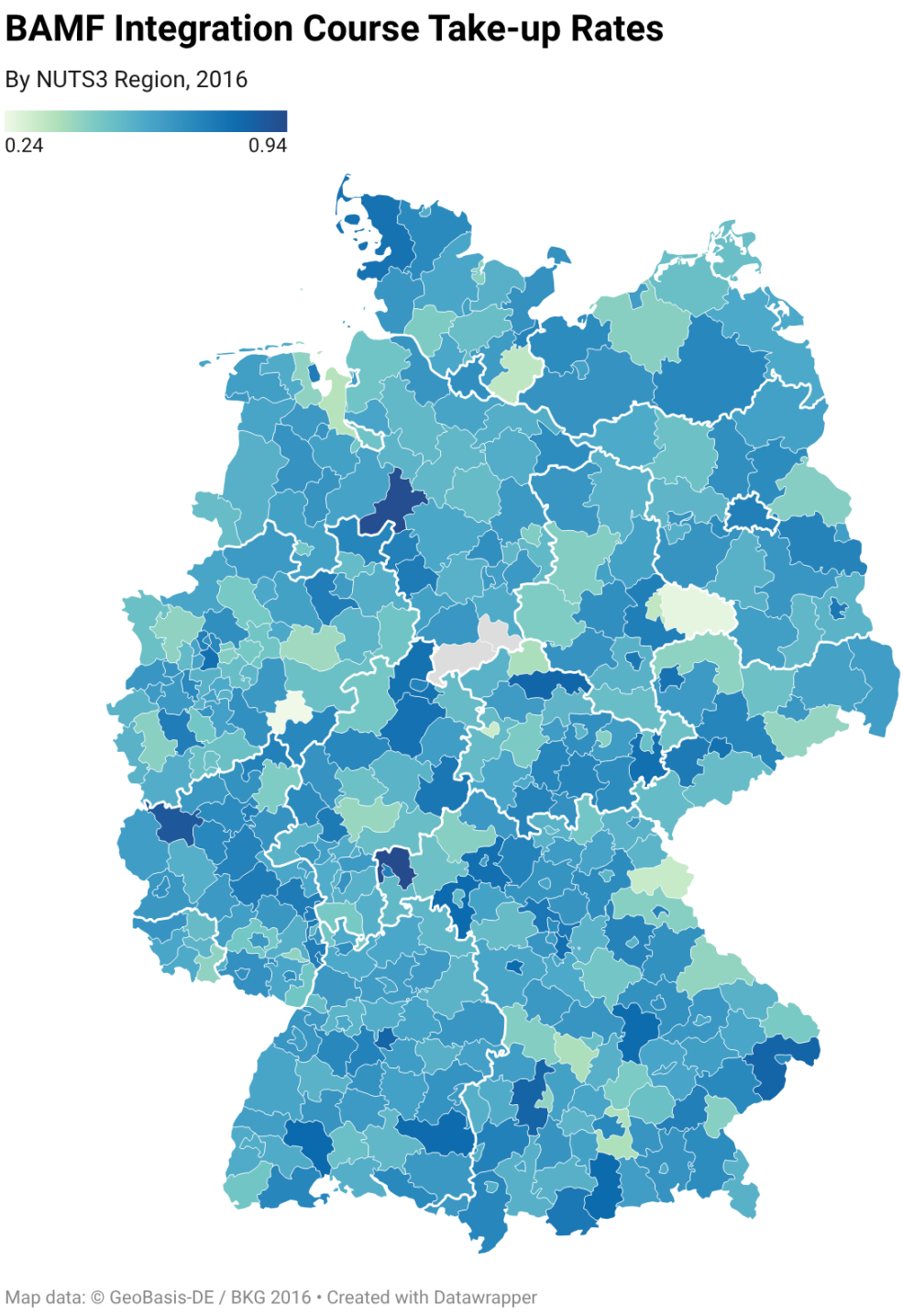
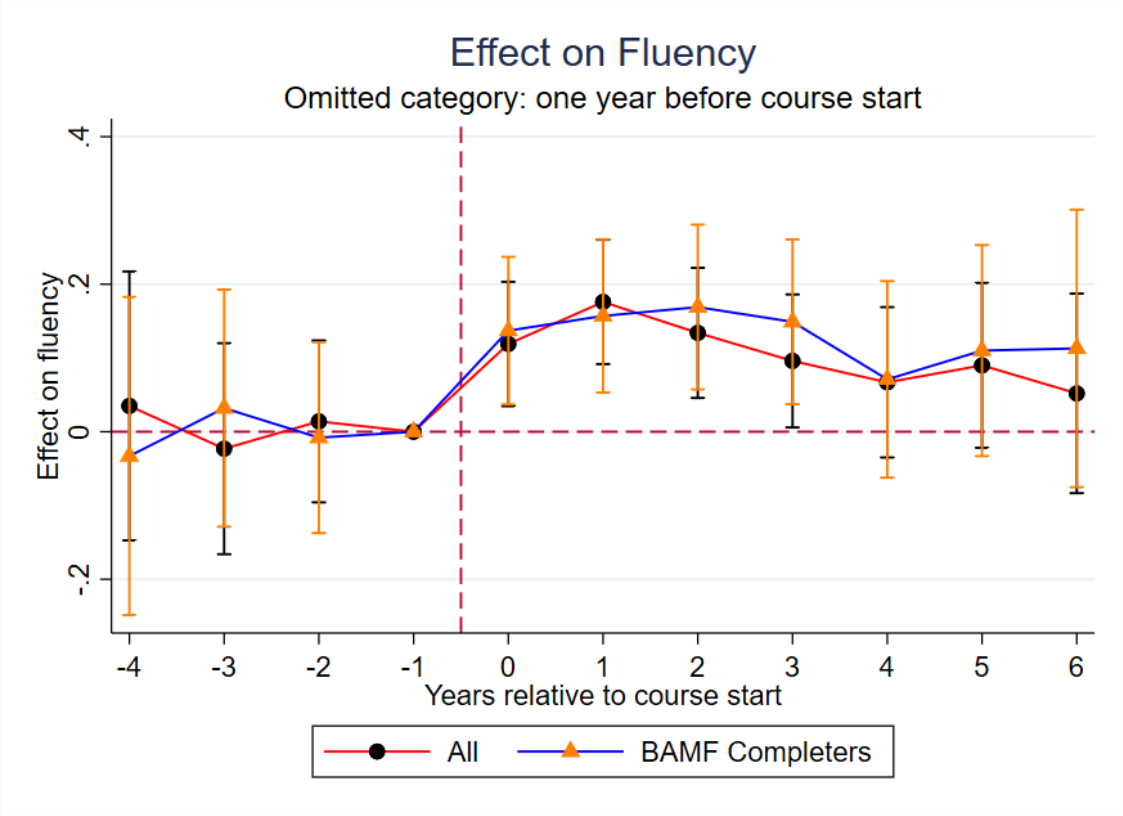


Figure 2: BSK Dynamic Effects



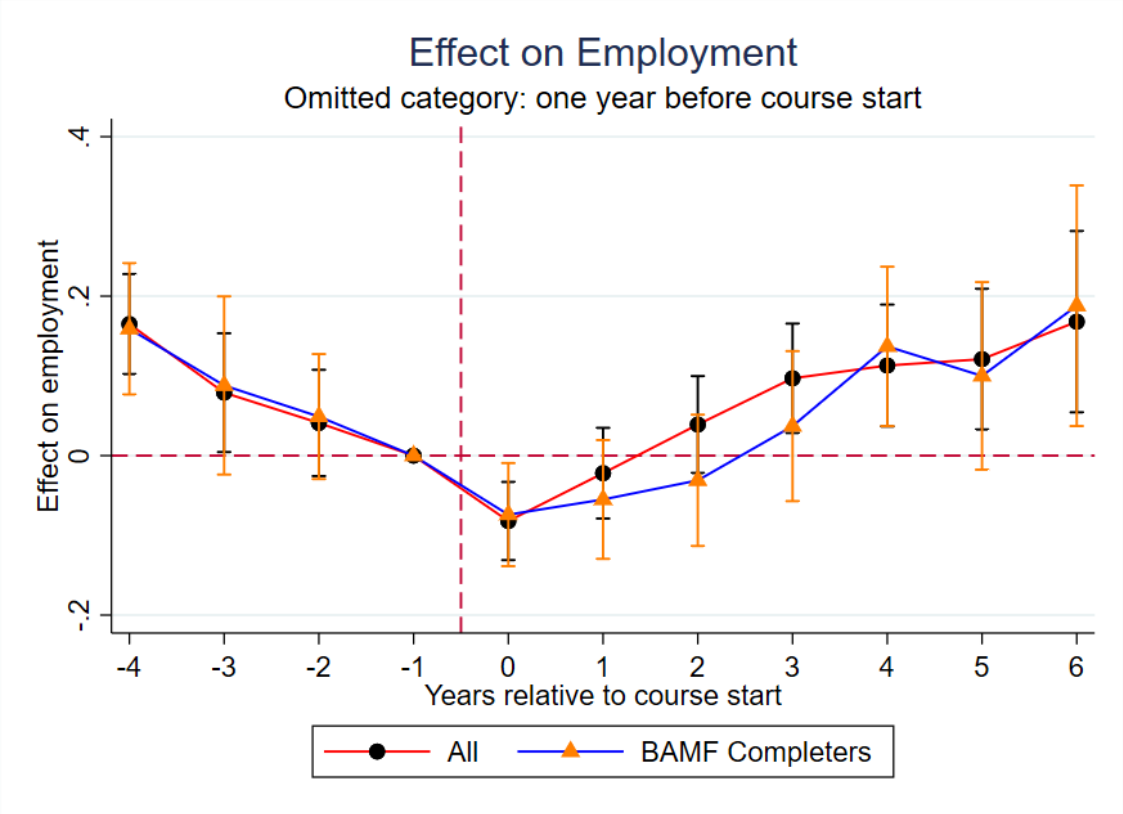
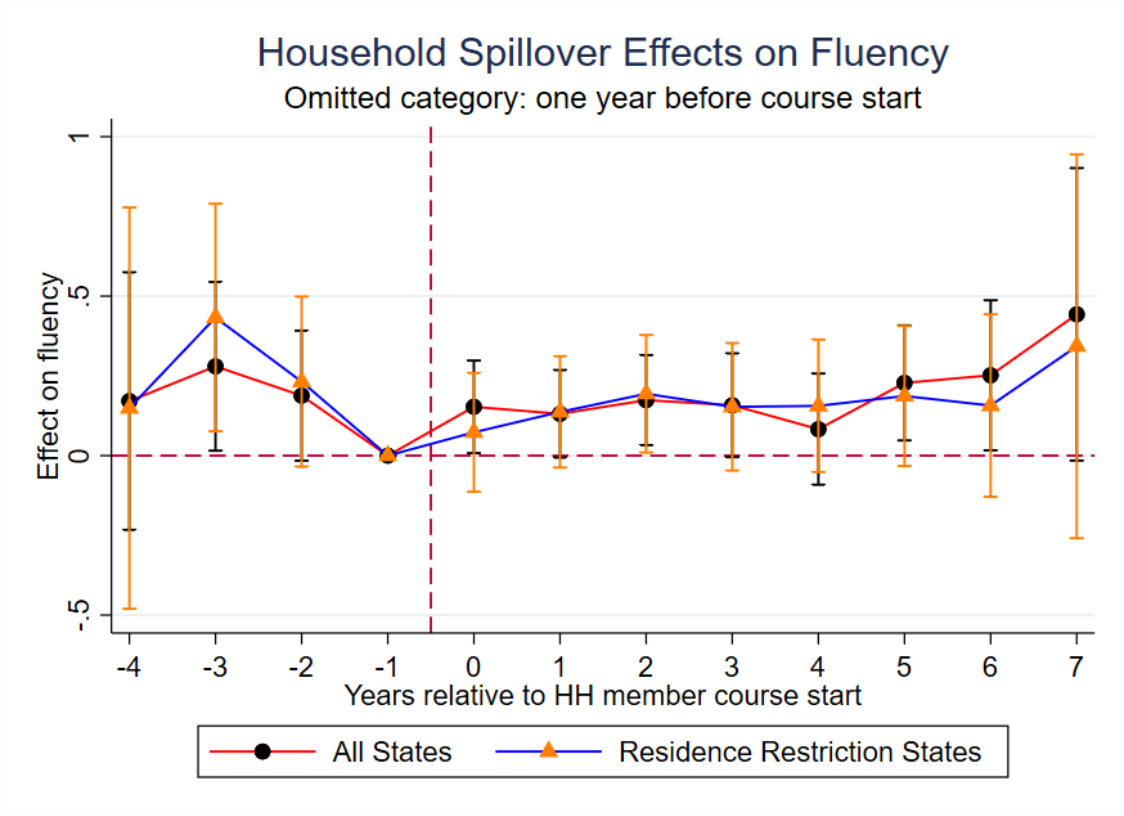


Figure 3: BSK Dynamic Household Effects



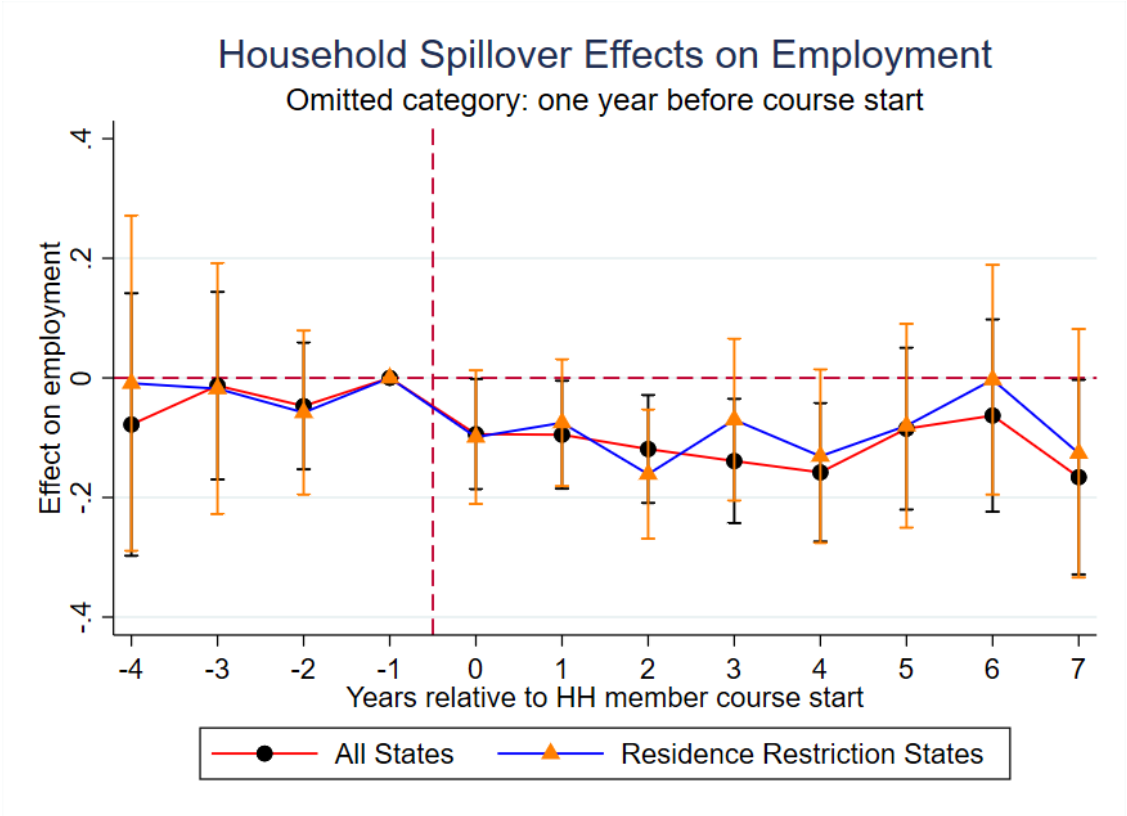
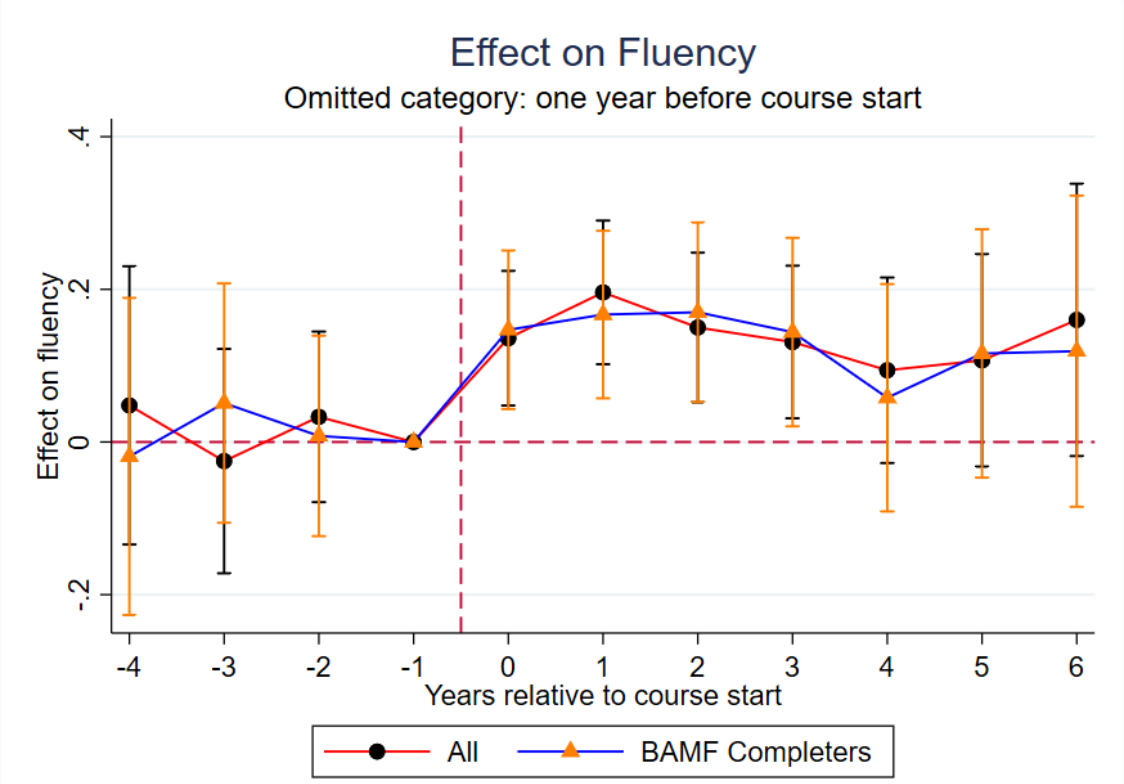
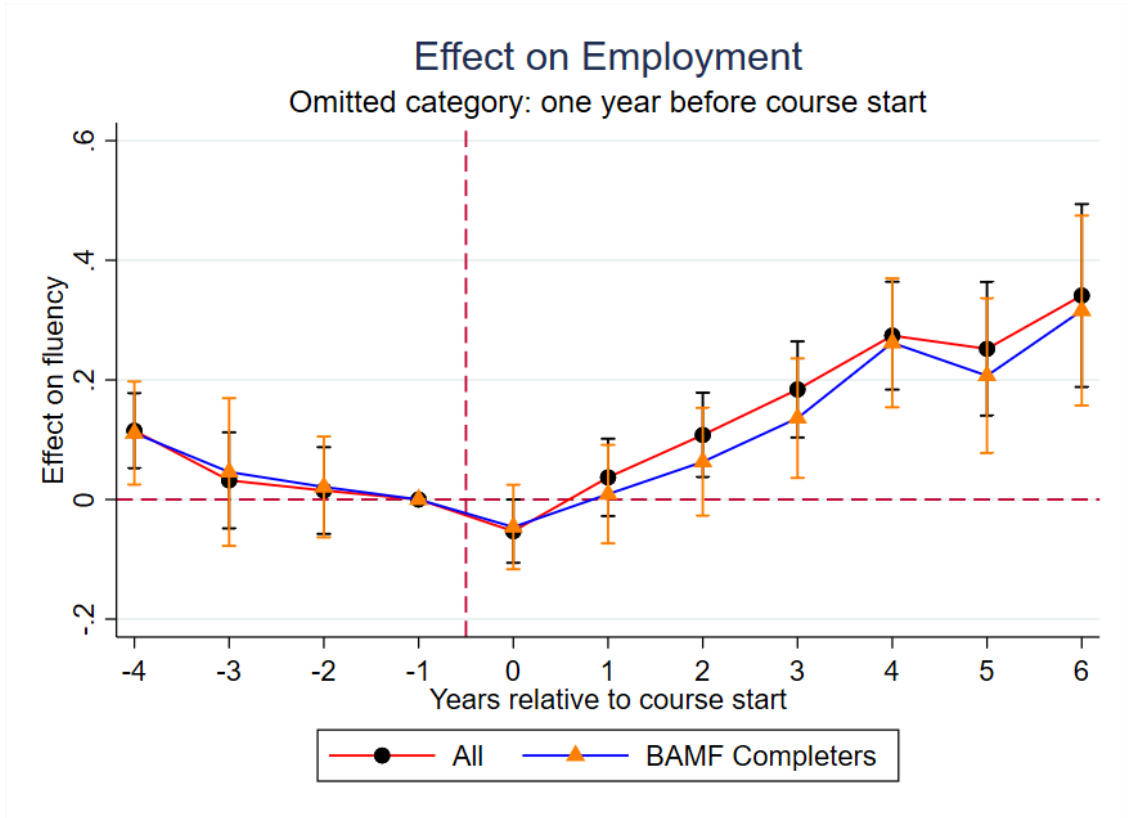


Figure 4: BSK Dynamic Effects, Overlap Weights





Appendix

Table A1: Descriptive Statistics by BSK Participation
Full Sample from Restrictive States

	Never Treated		Pre-BSK		Post-BSK	
Variable	Mean	Obs.	Mean	Obs.	Mean	Obs.
Fluency	-0.611	9,882	-0.680	557	-0.081	1,598
Employed	0.327	10,857	0.075	557	0.457	1,610
Hourly Wage	13.265	2,673	9.319	18	12.141	607
High-Skill Occupation	0.639	3,003	0.571	35	0.700	611
Female	0.390	10,857	0.357	557	0.267	1,610
BAMF Course Completion	0.574	10,857	0.722	557	0.906	1,610
Secondary Education	0.099	10,857	0.061	557	0.096	1,610
Post-Secondary Education	0.179	10,857	0.280	557	0.344	1,610
Years Since Migration	3.981	10,857	2.212	557	4.706	1,610
Married	0.593	10,857	0.659	557	0.580	1,610
# of Children in Household	1.667	10,857	1.864	557	1.487	1,610
Lagged Median Prejudice	0.021	10,857	0.026	557	0.019	1,610

Notes: This table includes only observations surveyed from 2015 to 2024 and immigrated from 2015 to 2018. Language fluency is standardized from questions on German reading, writing, and speaking ability. Both education variables, Married, and Female are binary. BAMF Course is also binary and indicates whether the BAMF integration course has been taken. Hourly wage is in 2012 Euro. High-Skill Occupation is equal to one if the respondent has a KLDB skill level of at least 2, indicating at least some vocational or post-secondary qualification. Lagged median prejudice calculated from the weighted distribution on immigration attitude questions nine years before the survey year.

Table A2: Descriptive Statistics by BSK Participation
Post-BAMF Course Only, Restrictive States

	Never Treated		Pre-BSK		Post-BSK	
Variable	Mean	Obs.	Mean	Obs.	Mean	Obs.
Fluency	-0.490	6,175	-0.521	402	-0.031	1,268
Employed	0.349	6,226	0.097	402	0.481	1,276
Hourly Wage	12.736	1,672	9.319	18	12.203	530
High-Skill Occupation	0.615	1,830	0.625	32	0.692	514
Female	0.342	6,226	0.331	402	0.263	1,276
Secondary Education	0.085	6,226	0.070	402	0.090	1,276
Post-Secondary Education	0.198	6,226	0.299	402	0.374	1,276
Years Since Migration	4.213	6,226	2.525	402	4.825	1,276
Married	0.624	6,226	0.662	402	0.602	1,276
# of Children in Household	1.668	6,226	1.789	402	1.527	1,276
Lagged Median Prejudice	0.023	6,226	0.026	402	0.023	1,276

Notes: This table includes only observations surveyed from 2015 to 2024 and immigrated from 2015 to 2018. Language fluency is standardized from questions on German reading, writing, and speaking ability. Both education variables, Married, and Female are binary. BAMF Course is also binary and indicates whether the BAMF integration course has been taken. Hourly wage is in 2012 Euro. High-Skill Occupation is equal to one if the respondent has a KLDB skill level of at least 2, indicating at least some vocational or post-secondary qualification. Lagged median prejudice calculated from the weighted distribution on immigration attitude questions nine years before the survey year.

Table A3: Stacked Event Study Estimates of BSK Participation

	All States				Residence Restriction States			
	All Refugees		BAMF Completers		All Refugees		BAMF Completers	
Event Time	Fluency	Employment	Fluency	Employment	Fluency	Employment	Fluency	Employment
-4	0.035	0.165***	-0.033	0.159***	0.147	0.195***	0.112	0.187***
	(0.093)	(0.032)	(0.110)	(0.042)	(0.126)	(0.039)	(0.151)	(0.057)
-3	-0.023	0.079**	0.032	0.088	0.015	0.104**	0.016	0.099
	(0.073)	(0.038)	(0.082)	(0.057)	(0.085)	(0.048)	(0.102)	(0.086)
-2	0.014	0.041	-0.008	0.049	0.013	0.060	-0.118	0.057
	(0.056)	(0.034)	(0.066)	(0.040)	(0.072)	(0.041)	(0.085)	(0.052)
0	0.119***	-0.082***	0.137***	-0.074**	0.156***	-0.084***	0.100	-0.083**
	(0.043)	(0.025)	(0.051)	(0.033)	(0.054)	(0.031)	(0.064)	(0.041)
1	0.176***	-0.022	0.157***	-0.055	0.243***	-0.020	0.138**	-0.063
	(0.043)	(0.029)	(0.053)	(0.038)	(0.054)	(0.035)	(0.066)	(0.046)
2	0.134***	0.039	0.169***	-0.031	0.213***	0.042	0.188***	-0.027
	(0.045)	(0.031)	(0.057)	(0.042)	(0.056)	(0.037)	(0.070)	(0.049)
3	0.096**	0.097***	0.149***	0.037	0.170***	0.148***	0.125*	0.088
	(0.046)	(0.035)	(0.057)	(0.048)	(0.058)	(0.043)	(0.072)	(0.056)
4	0.067	0.113***	0.071	0.137***	0.132**	0.183***	0.069	0.155**
	(0.052)	(0.039)	(0.068)	(0.051)	(0.065)	(0.047)	(0.083)	(0.062)
5	0.090	0.121***	0.110	0.100*	0.146**	0.111**	0.056	0.042
	(0.057)	(0.045)	(0.073)	(0.060)	(0.069)	(0.054)	(0.088)	(0.073)
6	0.052	0.168***	0.113	0.188**	0.106	0.203***	-0.034	0.155*
	(0.069)	(0.058)	(0.096)	(0.077)	(0.086)	(0.070)	(0.113)	(0.092)
Stack-Individual FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Stack-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Stacked Obs.	91,637	92,660	59,938	60,013	60,005	60,706	38,837	38,905

Notes: This table shows estimates from stacked event-study regressions with $t = -1$ as the omitted period. All specifications include stacked individual fixed effects, stack-year fixed effects, and the controls used in Table 4. Standard errors are clustered at the individual level and are reported in parentheses.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A4: Stacked Event Study:
Household Spillover Effects of BSK Course

	All States		Residence Restriction States	
Event Time	Fluency	Employment	Fluency	Employment
-4	0.171	-0.078	0.149	-0.009
	(0.206)	(0.112)	(0.321)	(0.143)
-3	0.280**	-0.013	0.433**	-0.018
	(0.135)	(0.080)	(0.182)	(0.107)
-2	0.188*	-0.047	0.232*	-0.058
	(0.104)	(0.054)	(0.136)	(0.070)
0	0.153**	-0.094**	0.073	-0.099*
	(0.074)	(0.047)	(0.095)	(0.057)
1	0.131*	-0.095**	0.137	-0.075
	(0.070)	(0.046)	(0.089)	(0.054)
2	0.174**	-0.119**	0.194**	-0.161***
	(0.072)	(0.046)	(0.094)	(0.055)
3	0.158*	-0.139***	0.153	-0.070
	(0.083)	(0.053)	(0.102)	(0.069)
4	0.083	-0.158***	0.156	-0.131*
	(0.089)	(0.059)	(0.106)	(0.074)
5	0.228**	-0.085	0.187*	-0.080
	(0.092)	(0.069)	(0.112)	(0.087)
6	0.252**	-0.063	0.157	-0.003
	(0.120)	(0.082)	(0.146)	(0.098)
7	0.443*	-0.166**	0.343	-0.126
	(0.234)	(0.083)	(0.307)	(0.106)
Individual FE	Yes	Yes	Yes	Yes
Stack-Year FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Stacked Obs.	80,128	80,977	52,650	53,227

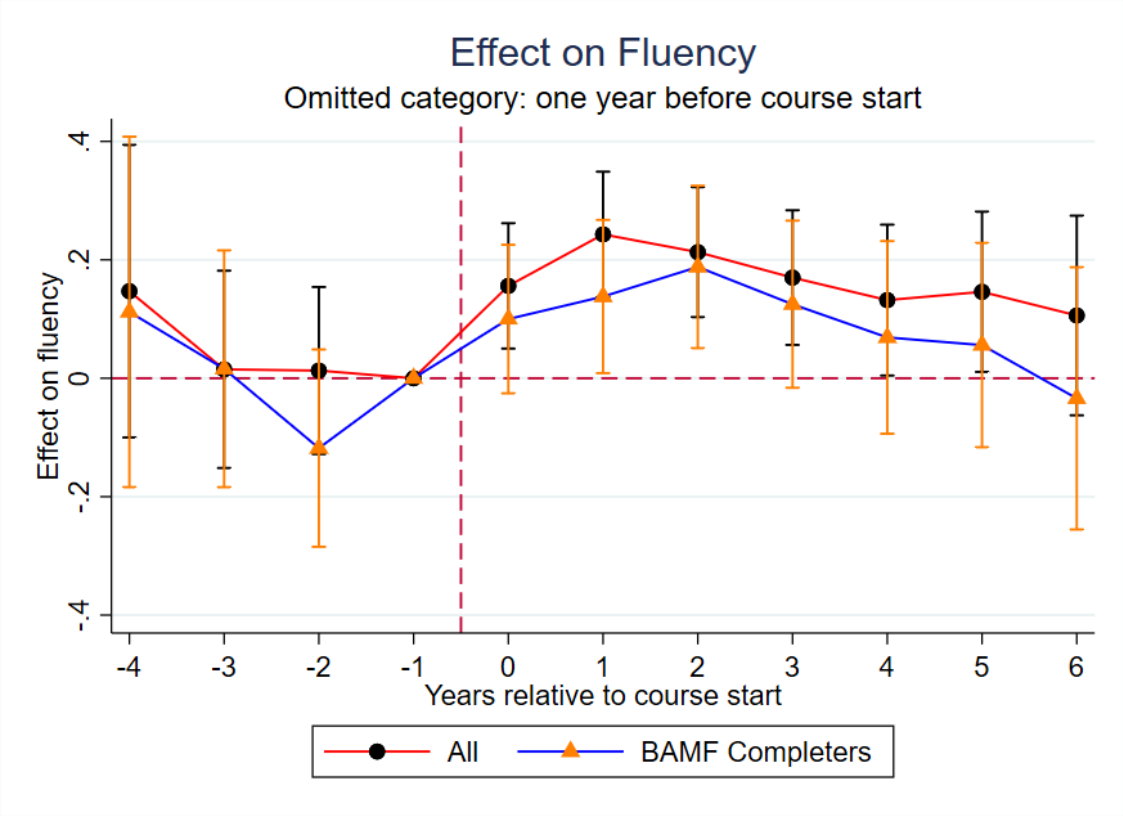
Notes: The dependent variables are the fluency and employment outcomes of other adult household members. Estimates are from stacked event-study regressions with $t = -1$ as the omitted category. All specifications include individual fixed effects, stack-year fixed effects, and the controls used in the main regressions. Standard errors are clustered at the household level and are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A5: Stacked Event Study Estimates of BSK Participation
Overlap Weighting

	All States				Residence Restriction States			
	All Refugees		BAMF Completers		All Refugees		BAMF Completers	
Event Time	Fluency	Employment	Fluency	Employment	Fluency	Employment	Fluency	Employment
-4	0.048	0.115***	-0.019	0.111**	0.154	0.123***	0.133	0.119*
	(0.093)	(0.032)	(0.106)	(0.044)	(0.129)	(0.043)	(0.148)	(0.068)
-3	-0.025	0.032	0.051	0.046	0.012	0.052	0.031	0.061
	(0.075)	(0.041)	(0.080)	(0.063)	(0.087)	(0.053)	(0.101)	(0.095)
-2	0.033	0.015	0.008	0.021	0.026	0.030	-0.110	0.015
	(0.057)	(0.037)	(0.067)	(0.043)	(0.075)	(0.042)	(0.089)	(0.054)
0	0.136***	-0.053*	0.147***	-0.046	0.165***	-0.060*	0.101	-0.060
	(0.045)	(0.027)	(0.053)	(0.036)	(0.056)	(0.033)	(0.067)	(0.044)
1	0.196***	0.037	0.167***	0.009	0.248***	0.057	0.140**	0.009
	(0.048)	(0.033)	(0.056)	(0.042)	(0.059)	(0.039)	(0.071)	(0.051)
2	0.150***	0.108***	0.170***	0.063	0.212***	0.117***	0.161**	0.078
	(0.050)	(0.036)	(0.060)	(0.046)	(0.061)	(0.040)	(0.075)	(0.052)
3	0.131**	0.184***	0.144**	0.136***	0.204***	0.256***	0.108	0.214***
	(0.051)	(0.041)	(0.063)	(0.051)	(0.065)	(0.046)	(0.079)	(0.059)
4	0.094	0.274***	0.058	0.262***	0.134*	0.341***	0.038	0.296***
	(0.062)	(0.046)	(0.076)	(0.055)	(0.075)	(0.053)	(0.091)	(0.064)
5	0.107	0.252***	0.116	0.207***	0.107	0.207***	0.045	0.166**
	(0.071)	(0.057)	(0.083)	(0.066)	(0.087)	(0.067)	(0.102)	(0.078)
6	0.160*	0.341***	0.119	0.316***	0.071	0.302***	-0.086	0.285***
	(0.091)	(0.078)	(0.104)	(0.081)	(0.105)	(0.094)	(0.123)	(0.098)
Stack-Individual FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Stack-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Stacked Obs.	68,324	69,089	45,400	45,464	44,727	45,236	29,419	29,477

Notes: This table shows estimates from overlap-weighted stacked event-study regressions with $t = -1$ as the omitted period. All specifications include stacked individual fixed effects, stack-year fixed effects, and the controls used in Table 4. Standard errors are clustered at the individual level and are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure A1: BSK Dynamic Effects, Residence Restriction States



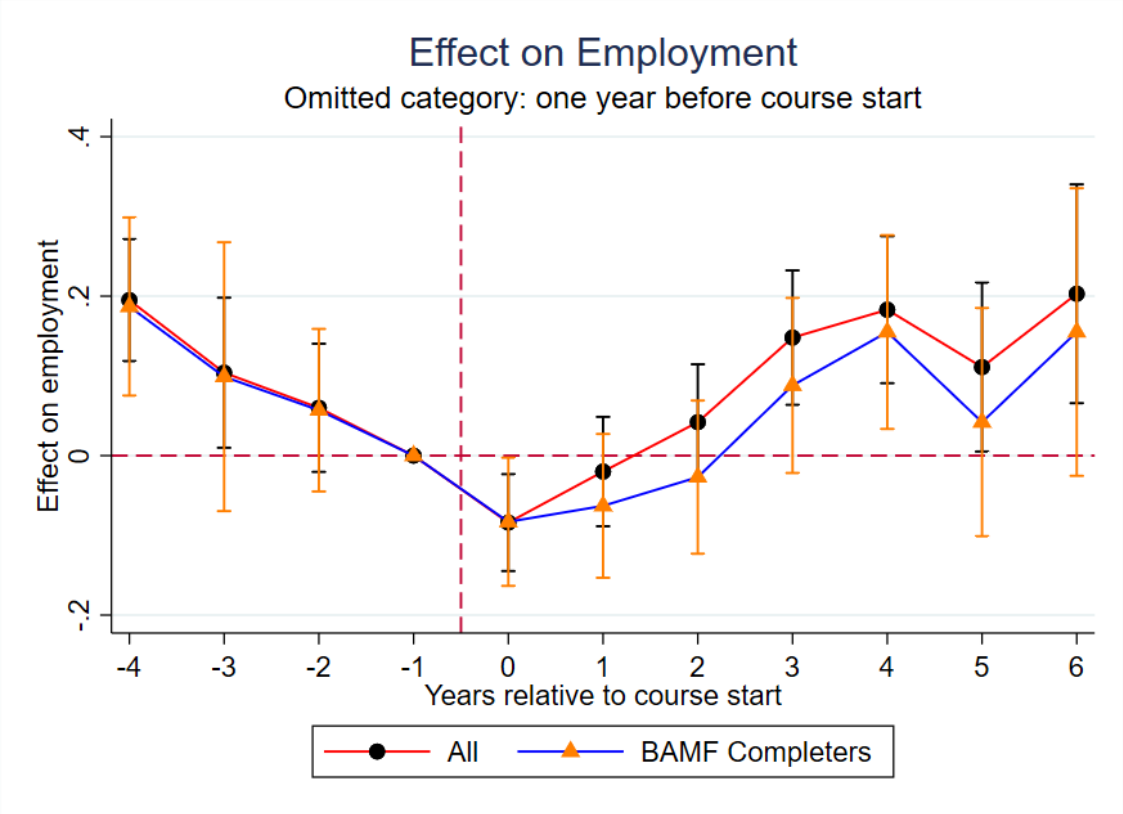
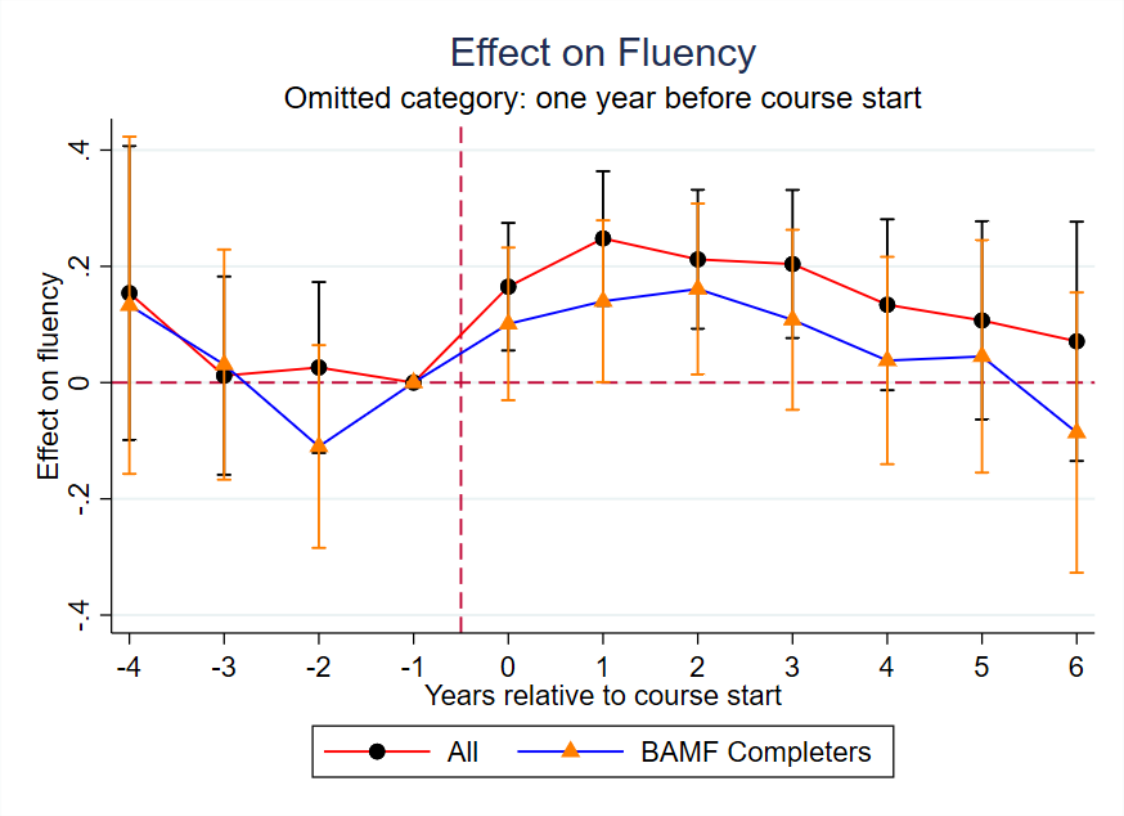
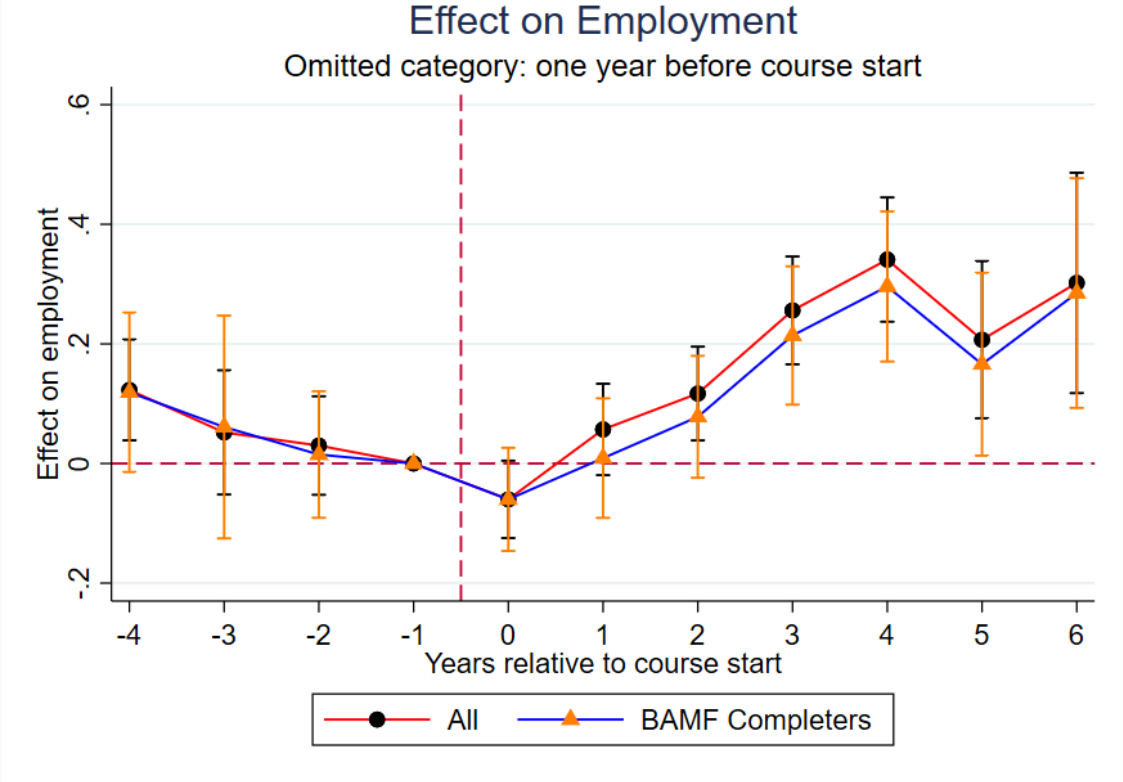


Figure A2: BSK Dynamic Effects, Overlap Weights, Residence Restriction States





References

Albarosa, E., and B. Elsner. "Forced Migration and Social Cohesion: Evidence from the 2015/16 Mass Inflow in Germany." World Development 167 (July 2023). https://doi.org/10.1016/j.worlddev.2023.106228

Arendt, Jacob Nielsen, Iben Bolvig, Mette Foged, Linea Hasager, and Giovanni Peri. "Language training and refugees' Integration." NBER Working Paper (2021). https://doi.org/10.3386/w26834.

Baderschneider, Ariane, Holger Bonin, Bernhard Boockmann, Marco Caliendo, Matthias Gensicke, Annabelle Krahn, and Steffen Zinsmeister. Evaluation der berufsbezogenen Deutschsprachförderung nach §45a Aufenthaltsgesetz. IZA Research Report No. 142 (2024). Bonn: Institute of Labor Economics (IZA). https://docs.iza.org/report_pdfs/iza_report_142.pdf.

BAMF. Integrationskursgeschäftsstatistik 2013-2019 (Landkreise und kreisfreie Städte). Integrationskursgzahlen; Federal Office for Migration and Refugees (BAMF) (excel files available upon request from the BAMF).

Brücker H, Rother N, Schupp J (eds.); Babka von Gostomski C, Böhm A, Brücker H, Fendel T, Friedrich M, Giesselmann M, Holst E, Kosyakova Y, Kroh M, Liebau E, Richter D, Romiti A, Rother N, Schacht D, Scheible JA, Schmelzer P, Schupp J, Siegert M, Sirries S, Trübswetter P, Vallizadeh E. "IAB BAMF-SOEP-Befragung von Geflüchteten: Überblick und erste Ergebnisse" IAB-Forschungsbericht 14 (2016), Nuremberg, Germany.

Cengiz, Doruk, Arindrajit Dube, Attila Lindner, Ben Zipperer. "The Effect of Minimum Wages on Low-Wage Jobs," The Quarterly Journal of Economics, President and Fellows of Harvard College, vol. 134(3) (August 2019). 1405-1454. https://doi.org/10.1093/qje/qjz014.

Chin, Aimee, and Kalena E. Cortés. "The Refugee/Asylum Seeker." In *Handbook of the Economics of International Migration*, vol. 1. Elsevier, 2015. https://doi.org/10.1016/B978-0-444-53764-5.00012-8.

Chiswick, Barry R., and Paul W. Miller. "The Endogeneity between Language and Earnings: International Analyses." *Journal of Labor Economics* 13, no. 2 (1995): 246-88. https://doi.org/10.1086/298374.

Cumming, Philippa, and Ellen Heidinger. "Refugees Are Feeling Increasingly Unwelcome and Perceiving Discrimination in the Labor and Housing Markets." *DIW Weekly Report*, ahead of print, 2025. https://doi.org/10.18723/DIW_DWR:2025-34-1.

De Vroome, Thomas, and Frank Van Tubergen. "The Employment Experience of Refugees in the Netherlands." International Migration Review 44, no. 2 (2010): 376–403. https://doi.org/10.1111/j.1747-7379.2010.00810.x.

Dustmann, Christian, and Arthur Van Soest. "Language and the Earnings of Immigrants." *Industrial and Labor Relations Review* 55, no. 3 (2002): 473.

https://doi.org/10.2307/2696052.

Jaschke, Philipp, Sulin Sardoschau, and Marco Tabellini. "Scared Straight? Threat and Assimilation of Refugees in Germany." NBER Working Paper No. 30381 (2022). https://doi.org/10.3386/w30381.

Kanas, Agnieszka, and Yuliya Kosyakova. "Greater Local Supply of Language Courses Improves Refugees' Labor Market Integration." European Societies 25, no. 1 (2023): 1–36. https://doi.org/10.1080/14616696.2022.2096915.

Kaida, Lisa. "Do Host Country Education and Language Training Help Recent Immigrants Exit Poverty?" Social Science Research 42, no. 3 (2013): 726–41. https://doi.org/10.1016/j.ssresearch.2013.01.004.

Katz, Bruce, Luise Noring, and Nantke Garrelts. "Cities and Refugees: The German Experience." Brookings Institution (2016). https://www.brookings.edu/articles/cities-and-refugees-the-german-experience/.

Krause-Pilatus, A., Rinne, U. "Wirksamkeit und Effizienz der arbeitsmarktpolitischen Integrationsmaßnahmen: Ergebnisse und Handlungsempfehlungen für Geflüchtete in Deutschland." Sozialer Fortschritt Vol.72, Iss. 1 (2023). 23-41.

Kugler, Adriana, Maurice Kugler, Juan E. Saavedra, and Luis Omar Herrera-Prada. "Long-Term Educational Consequences of Vocational Training in Colombia: Impacts on Young Trainees and Their Relatives." Journal of Human Resources 57, no. 1 (2022): 178–216. https://doi.org/10.3368/jhr.57.1.0518-9528R2.

Li, Fan, and Laine E. Thomas. "Addressing Extreme Propensity Scores via the Overlap Weights." American Journal of Epidemiology, ahead of print, September 5, 2018. https://doi.org/10.1093/aje/kwy201.

Lochmann, Alexia, Hillel Rapoport, and Biagio Speciale. "The Effect of Language Training on Immigrants' Economic Integration: Empirical Evidence from France." European Economic Review 113 (April 2019): 265–96. https://doi.org/10.1016/j.euroecorev.2019.01.008.

Oreopoulos, Philip. "Why Do Skilled Immigrants Struggle in the Labor Market? A Field Experiment with Thirteen Thousand Resumes." *American Economic Journal: Economic Policy* 3, no. 4 (2011): 148-71. https://doi.org/10.1257/pol.3.4.148.

Ruiz, Isabel, and Carlos Vargas-Silva. "Differences in Labour Market Outcomes between Natives, Refugees and Other Migrants in the UK." *Journal of Economic Geography* 18, no. 4 (2018): 855-85. https://doi.org/10.1093/jeg/lby027.

Runst, Petrik. "The Effect of Occupational Licensing Deregulation on Migrants in the German Skilled Crafts Sector." *European Journal of Law and Economics* 45, no. 3 (2018): 555-89. https://doi.org/10.1007/s10657-018-9583-x.

Schlosser, Analia, and Yannay Shanan. "Fostering Soft Skills in Active Labor Market Programs: Evidence from a Large-Scale RCT." Journal of Human Resources (January 8, 2025). https://doi.org/10.3368/jhr.0723-13013R1.

Titus, David. "Market Structures, Prejudice, and the Residual Wage Gap Between Refugees and Natives." Working Paper (2026).

Waisman, Gisela, and Birthe Larsen. "Do Attitudes Towards Immigrants Matter?" *SULCIS Working Papers* (2008).

Yao, Yuxin, and Jan C. Van Ours. "Language Skills and Labor Market Performance of Immigrants in the Netherlands." *Labour Economics* 34 (June 2015): 76-85. https://doi.org/10.1016/j.labeco.2015.03.005.