Root Growing and Path Dependence in Location Choice: Evidence from Danish Refugee Placement

Farid Farrokhi
Purdue

David Jinkins*
Copenhagen Business School

December 15, 2023

Abstract

Does spending time in a location *cause* a person to stay there longer? We use a 1999 change in Danish refugee settlement policy to address this question. The policy change strongly encouraged refugees to stay in their assigned settlement municipality for at least three years. Using empirical designs for natural experiments, we find that treated refugees were more likely to be in their assigned location many years after their residence was granted. In a difference-in-differences specification, treated refugees were 4.8 percentage points more likely to remain in their first commuting zone 13 years later. A regression discontinuity design delivers a larger but less precise point estimate.

Keywords: Location choice, Path dependence, Moving costs
JEL Classification: I38, J61, R23

1 Introduction

A long-standing tradition in economic geography has emphasized the importance of path dependence in the location of factors of production (Krugman, 1991; Bleakley and Lin, 2012). While it is well-documented that cities are persistent as the locations of economic activity, the specific mechanisms that cause the persistence are still unknown. We take a step toward identifying such mechanisms by studying the causal effect of staying in a location on an individual's likelihood of staying there longer. We call this phenomenon "root

*This paper was formerly circulated under the title "Dynamic Amenities and Path-Dependence in Location Choice". We are grateful for funding from the Independent Research Fund Denmark grant number 8019-00031B which made this project possible. We would also like to thank for helpful comments Niels Henning Bjørn, Anna Piil Damm, Laurent Gobillon, David Hummels, Birthe Larson, Rie Martens, Marie Louise Schultz-Nielsen, Herdis Steingrimsdottir, Chong Xiang, participants at a NOITS meeting, a Meeting of the Urban Economics Association, a Kraks Fund seminar, and seminars at Copenhagen Business School and Purdue, as well as two anonymous referees. All remaining errors are our own.

growing." In contrast to conventional amenities that are independent of a person's history of residential location, a person's roots grow over time in her particular location. The benefit that a person derives from living in a city rises as the person learns about local amenities, as she gets attached to them, and as her social network expands.

Many people have lived where they are for many years and are attached to their home city. We can confirm in our Danish data that the longer someone has been in a location, the less likely she is to leave. This relationship is robust to controlling for marital status, age, number of children, and home ownership status. But that is not enough to show what we are after. To take an example from the US, the people who remain in Washington DC for several years are likely those who do not mind a long commute or a humid summer. The long-term residents of a city are *selected* into being unlikely to leave, for reasons that have nothing to do with deepening roots. To separate root growing from such underlying correlations, we examine a counterfactual question—namely, if a person who left a city had spent a few more years there, would she have been less likely to leave?

It is challenging to separate the selection of long-term residents from the causal effect of spending time in a city. To address this challenge, we use variation from a policy change in the placement of refugees in Denmark. Prior to 1999, refugees arriving in Denmark were allocated housing in different Danish municipalities. They were encouraged by social workers to stay in their placement municipality for at least eighteen months. After a law change in 1999, refugees were strongly incentivized to remain in their placement municipality for three years.[1] A literature on Danish refugee settlement has argued that refugee placements were random conditional on several factors such as family status and country of origin (Damm, 2005; Azlor et al., 2020). Since most of these factors are observable and can be controlled for, we follow the literature in treating placements as conditionally random.

If spending time in a location causes people to grow roots, then refugees treated by the policy change will be more likely to be observed in their placement commuting zone many years after their arrival. To test this hypothesis, we use methods borrowed from the causal inference literature, namely regression discontinuity (RD) and difference-in-differences (DiD) designs, as well as a duration-model analysis. In the RD and DiD designs, the outcome variable is whether an immigrant is still in her first location thirteen years after being granted a visa in Denmark. The longest required stay is three years, so this is ten years after the end of the required stay. In the duration model, we trace out how a refugee's hazard rate of leaving and survival rate of remaining in their commuting zone of the initially allocated municipality depends on whether he arrived before or after the policy change.

[1]The refugee placement policy both before and after 1999 involved many subtleties, and the 1999 reform had other aspects in addition to the lengthening and strengthening of the required residence period. We discuss the placement policy and related reforms in detail in Section 2.

All three designs provide evidence that refugees who were treated were more likely to stay in their placement commuting zone. According to the RD design, post-reform refugees relative to pre-reform refugees were 11.3 percentage points more likely to stay in their initial placement 13 years after the residence was granted, although the estimate is not significant at a conventional five-percent significance level. According to the DiD design—with non-refugee immigrants from the same origin country placed in the same initial commuting zone as the control group—treated refugees were 4.8 percentage points more likely to remain in their placement 13 years later.[2] According to the duration model—which uses the same control group as in our DiD desig—a higher fraction of treated refugees, equal to 2.4 percentage points, were expected to remain in their initial commuting zone 13 years after placement. These complementary results support our root growing hypothesis that spending time in a location increases the value of a longer stay.

We explore mechanisms behind these results by replacing the outcome variable in our difference-in-differences specification with other characteristics of refugees. We find that, after the policy, refugee household heads were more likely to have children in their first years in Denmark and were more likely to be working. The policy appeared to have no effect on marriage status or on the incidence of receiving government support. These results suggest that being forced to stay in a location made refugees more likely to exert effort to look for a local job, and have children earlier. It may be that deeper social connections to coworkers, other parents, and other children may be the tie that binds treated residents more strongly to their settlement locations.

This paper contributes to the urban economics and economic geography literature in several ways. Tombe and Zhu (2019) find that migrating between China's provinces reduces lifetime utility on average by nearly a factor of three. Kennan and Walker (2011) estimate that average moving costs for movers between American states amount to $312,000 (in 2010 dollars). These large estimates point to a range of migration frictions above and beyond direct, monetary moving costs.[3]

More generally, the literature has shown, in a number of different contexts, the importance of history in the current location of economic activity (Bleakley and Lin, 2012; Kline and Moretti, 2013; Dalgaard et al., 2018).[4] Our causal findings relating tenure to location choice complement studies that emphasize

[2]The control group includes all non-refugee immigrants except students and those on family reunification visas.

[3]Additionally, a number of studies have shown that home ownership reduces mobility (Munch et al., 2006; Mian and Sufi, 2014; Yagan, 2014), and that people tend to remain in their locations even when the location is hit by a negative productivity shock (Autor et al., 2015).

[4]The literature suggests that the importance of history in the current location of cities may depend on the empirical context. Davis and Weinstein (2002) examine the persistence of the distribution of population in Japan, finding that the long-run city size in Japan was robust to shocks as big as the destruction of Japanese cities in WWII. Miguel and Roland (2011) also find support for persistence when considering the US bombing of Vietnam as a shock. Bosker et al. (2007), however, find support for multiple equilibria in German city growth given the bombing of Germany in WWII. Likewise, Michaels and Rauch (2018) find support for the importance of history by examining the effect of the collapse of the Western Roman Empire on city locations in Britain and France.

the importance of local ties or home attachment in the tendency of a person to reside near his birthplace (Zabek, 2019; Coate and Mangum, 2019). On the theoretical side, Allen and Donaldson (2022) study the path of spatial equilibrium and its potential dependence on history. This article contributes to the empirical part of this literature by showing that part of the persistence in spatial distribution of economic activity is driven by individuals growing roots where they live.

Our empirical study is inspired by an extensive literature using natural experiments in urban and labor economics (Baum-Snow and Ferreira, 2015). Several studies have taken advantage of forced migration to study the benefits of moving on future human capital and earning (Nakamura et al., 2016; Voigtlaender et al., 2020). Most closely related to our study, we build on a literature that investigates the conditionally random allocation of refugees to cities in Nordic countries (Edin et al., 2003; Damm, 2005; Damm and Dustmann, 2014; Foged and Peri, 2016; Dustmann et al., 2018; Azlor et al., 2020). This literature often uses data from refugees or immigrants as instruments to study economic or social outcomes for the entire sample of society (e.g. How large are agglomeration economies? How does the number of immigrants in a location affect voter sentiment?) Focusing on refugees and immigrants themselves, Eckert et al. (2022) and Hybel et al. (2023) study the effect of big city experience on wages. Finally, Nielsen and Jensen (2006) document many of the same data trends we do for the 1999 policy reform, albeit over a shorter period.

2 Policy Background

In this section, we describe the Danish policy for settling refugees in the period up to 1998, and changes made in 1999.

2.1 Danish Refugee Policy 1986-1998

Starting in the 1980's and up until 1999, the practicalities of assigning refugees to permanent housing in particular locations was handled by the non-profit Danish Refugee Council (DRC). As in other Scandinavian countries, low-cost housing is offered to asylum seekers who are granted refugee status in Denmark.[5] Our description of the policy in this period summarizes Damm (2005), and the interested reader should read that paper for more detail. In the early 1980's, refugees were assigned housing in their preferred location. Due to difficulty finding housing, and also because of political pressure to disperse refugees out of major urban

[5]Sometimes in this literature, the housing will be described as permanent. In Denmark, this often means a rental contract without a time-limit. Due to renter protection laws, it is difficult to evict renters as long as they pay their rent. A typical rental contract in the spot market has a fixed term, allowing landlords to evict the tenant at the end of the contract.

centers, in 1986 the DRC started a policy of conditionally random allocation of refugees to counties.[6]

The goal of the policy was to distribute refugees proportionally to the population of the 15 counties in Denmark. Only 182 of the 275 Danish municipalities received refugees in this period. The DRC only assigned refugees to municipalities which had necessary facilities for integration, which in practice meant that refugees were not usually placed in the most rural municipalities. The DRC had temporary offices in municipalities. The location of the offices would rotate over time across municipalities. Refugees were placed in the municipalities which had temporary offices when their refugee visas were granted. This mechanism led to clustering of nationalities in particular municipalities. Refugees did not need to accept the DRC housing offer if they could find their own permanent housing. In practice, around 90% of refugees took the housing offer from the DRC in the period 1986-1998. Because of a large inflow of refugees from the wars in the former Yugoslavia in the early 1990's, housing availability became the primary driver of placement and refugees were placed in rural municipalities as well.

An asylum seeker granted refugee status in this period was asked whether he would like help from the DRC in finding housing. If he agreed, he filled out a form including his preference of county and the reasons for that preference. Around ten days later, he would be provided temporary housing in a county, and the local office of the DRC would begin the search for permanent housing in the county. A refugee could ask to be reassigned if he was unhappy with the county he was allocated, and the DRC would reallocate him to another county. In practice this did not happen often. The average wait for permanent housing in the assigned county was 6-7 months. If the number of refugees requesting a particular county exceeded the quota for that county, priority was given to those with close family in the county or with special medical or educational needs available there. As refugee inflows to Denmark increased markedly in the 1990's, quotas in desirable counties filled up quickly in that period.

The literature that has studied Danish refugees' settlement policy indicates that the refugees' allocation across counties was random conditional on several controls. In particular, there were seven factors which influenced the initial assignment of refugees to locations: marital/family status, nationality, year of refugee status, location of close family or friends, special health needs, special educational needs, and finally a strong preference for a particular location. Of these, the first three are recorded in our data. Damm (2005) suggests proxies available in our data for the next three factors, and argues that the final factor is not too important since so few refugees asked to be reassigned. In an empirical analysis, she finds that all of the listed factors are statistically significant in influencing initial location assignment in the Danish register data, but combined their explanatory power is modest with an R-squared of approximately 0.13, which can be taken as suggestive evidence that refugee placement was indeed conditionally random.

[6]Prior to a reform in 2007, there were 14 Danish counties, and two "municipal" counties Copenhagen and Frederiksberg, both part of the Copenhagen metropolitan area. At the time these counties were divided up further into 275 municipalities.

After being placed in a county, the DRC provided instruction in Danish language, culture, and job training, and the refugees received a means-tested subsidy. Refugees were encouraged by social workers to stay in their assigned county for at least 18 months to complete the coursework, but they were free to leave as long as they could find housing elsewhere. Refugees could still receive the subsidy if they chose to leave.

2.2 The 1999 policy change

Partly in response to the large inflow of refugees from the former Yugoslavia in the mid 1990's, the Danish Integration Act was passed by the Danish legislature in 1998, coming into effect in 1999 (Udlaendinge- og Integrationsministeriet, 1998). Our understanding of the effects of this law change draws on the summary in Azlor et al. (2020), as well as our own reading of the law and interviews with Danish government bureaucrats who were involved with implementing the law. There were three important changes to the law that are relevant for our study. The first effect of the 1999 Danish Integration Act was that refugees became more tied to their assigned location. In particular, Azlor et al. (2020) report that the subsidy a refugee received from the government was conditional on the refugee residing in the assigned location for the extent of the training period. There was an exception to this rule if a refugee found full-time work in another municipality.[7] The second relevant effect of the 1999 reform was that the integration coursework was lengthened from 18 months to three years. Finally, the method for dispersing refugees across Danish counties changed.

With the passage of the new law, much of the responsibility for refugee settlement formerly delegated to the DRC was brought into the Danish government. The choice of municipalities to settle refugees was taken over by the Danish Immigration Service (DIS). Each year, DIS was tasked with estimating the number of refugees which would arrive the next year. These expected refugees were then allocated annually to municipalities based on the number of foreigners currently residing in the municipality and the total population of the municipality. If more refugees arrived than expected, DIS would adjust the quotas as needed. Locations with fewer foreigners relative to population were assigned higher refugee quota. Refugees would only rarely in special circumstances be allocated to a location with its quota already filled.[8] The placement

[7] We have obtained a sample of the letter sent to refugees when they received their visa from the Danish Immigration Service. This letter includes the following paragraph, loosely translated from Danish:

> You cannot choose where you will live yourself, because it is a large–and expensive–task for the municipality to organize an integration program. As a starting point, for you to follow an integration program and ultimately receive a permanent residence, it is a requirement that you live for 3 years in the municipality which the Ministry for Foreigners has decided you should live in. There can also be consequences for the payment of your subsidy if you insist on moving. You can find another municipality to live in if you have found permanent full-time work there.

In our interviews, one bureaucrat involved with refugee placement at the time of the reform said that while refugees believed they would lose their subsidy if they moved, it was not always true in practice.

[8] These parts of the Danish Integration Act of 1999 were largely unchanged until after 2012, when major changes to the law were passed by the Danish legislature particularly in 2016. These changes were partly a response to the then ongoing refugee

of a particular refugee was decided by DIS based in part on an interview with the refugee himself. In making the allocation decision, DIS considered factors such as whether the refugee had family already in Denmark, the nationality of the refugee, and any special educational or medical needs.

3 Data and Descriptive Statistics

Our empirical analysis relies on Danish register data. Because our primary empirical exercise focuses on an immigration policy change in 1999, we will mostly be interested in behavior of people who arrived in Denmark around the time of the policy change. Concretely, our data is composed of all working-age household heads who were granted a residence permit in Denmark from 1986 to 2016. We drop anyone who has close family connections in Denmark at the time of arrival. We break this sample of immigrants into two groups. The first is household heads granted refugee permits ("refugees"), and the second is household heads granted non-refugee residence permits ("non-refugee immigrants"). Since our paper is about the decision to stay in a location, the definition of location will be important. The municipality level in Denmark is too fine. Moving between the small municipalities often does not require changing job, school, or social group. Instead we use 21 commuting zones for Denmark based on those constructed by Eckert et al. (2022).[9]

3.1 Descriptive Statistics

We tabulate our descriptive statistics for all adult household heads separately for our two groups of Refugees and Non-refugee immigrants. Except where obvious, we limit our descriptive statistics to immigrants arriving in the years 1986-2003, because this corresponds to the sample in our empirical exercises below. We report statistics separately for immigrants arriving before and after the 1999 refugee policy change. All statistics we report correspond to the immigrants' first observed year in our data.

Figure 1 contains the number of immigrants in our sample population by year. The number of non-refugee immigrants before 1997 should be viewed cautiously, because the method we use classifies people less likely to be refugees (say from less common refugee sending countries) as a non-refugee immigrant, even if these people are in fact refugees. In particular, the drop in the level of non-refugee immigration from 1996 to 1997 is likely to simply be that we capture non-refugee immigrants more accurately in 1997.[10] The rapid increase in immigration after 2003 was due to the EU integration, which allowed free immigration

crisis in Europe induced by the Arab Spring and Syrian Civil War (Dustmann et al., 2017). In addition to generally becoming more strict on immigration, the rules about how long refugees need to remain in the allocated municipality and what types of training programs they must attend have been overhauled several times since 2012.

[9] Further details on how we cut and cleaned the data are in Appendix Section A.

[10] In Appendix B we show how our difference-in-difference estimates are affected by truncating the data from 1997 when our measure of refugee status becomes more precise.

from Eastern Europe. Growth in overall immigration stalled following the financial crisis of 2008, and then resumed with the recovery beginning in 2012. One somewhat startling observation to come out of this table is that while immigration has been rapidly growing over the period, Denmark granted markedly less refugee visas in the mid 2000's relative to the 1980's and 1990's. This trend was reversed during the Syrian refugee crisis in the 2010's.

The composition of refugee countries of origin changed over this period, depending on the location of wars and natural disasters in the world. Syrians and Afghanis were more likely to arrive after 1999, while Bosnians and Sri Lankans were more likely to arrive before 1999. Iran and Iraq were major refugee sending countries during both periods. We present a more detailed breakdown of the origin of refugees in our sample period in Figure A.4 in the appendix.

Table 1 contains descriptives for immigrant household heads. Refugee household heads are more likely to be men, while non-refugee household heads are slightly more likely to be women.[11] Refugees are more likely to be married when their permit was granted, and have both more children and larger family sizes. Non-refugees are younger on average than refugees particularly after 1999, because many of them arrive as students.

Table 2 presents the residence location of immigrant household heads in the first year after receiving their visas.[12] The most striking pattern visible in the table is that more than half of non-refugee immigrants live in Copenhagen when their visas are granted, but only a quarter of refugees live there. Refugees are also significantly more likely to live in the relatively rural and remote parts of Denmark outside of Central and South Jutland (Aarhus) and Fyn when their visas are granted.

Table 3 contains counts of visa types for non-refugee household heads after 1997 when our visa information starts. The two most popular forms of visa are for students and for general EU citizens. If we focused on non-household heads, family reunification would probably be more important. As it stands, this type of visa is rather unimportant, which is reassuring since we are worried about family relations of refugees falling into our control group. Student visas are a rather large category, and students often stay only for a few years and then leave their location within Denmark, or Denmark altogether. Since this profile is quite different from the standard refugee profile, we sometimes exclude student visas in our analysis below.

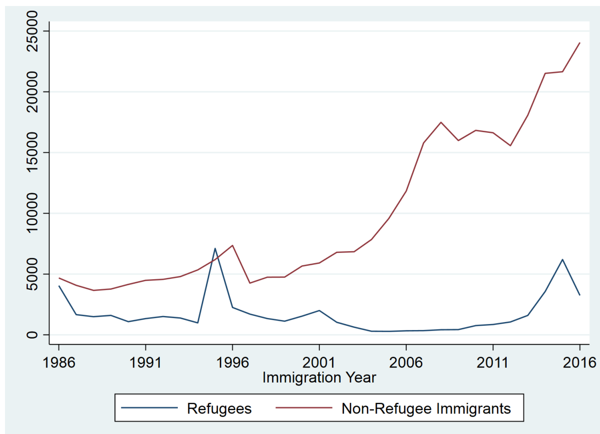
Finally, Table 4 reports unconditional staying probabilities for the sample of immigrant household heads. Refugee household heads are much more likely to remain in their placement commuting zones than non-

[11]If a married heterosexual couple is granted residence in the same year, we follow the Denmark Statistics protocol that the female is designated household head. If a married homosexual couple is granted residence, the protocol is that the older spouse is the household head.

[12]This table is based on somewhat larger Danish geographical regions rather than commuting zones due to disclosure rules.

refugee immigrants in all years after their initial placement.[13] Anticipating one of our estimation methods below, we also include the difference-in-differences for staying rates in the final column. While the numbers here are consistent with our root growing hypothesis, we should be careful not to interpret this table causally. As argued above, the assignment of refugees to locations can only be regarded as random conditional on the origin country and placement location of refugees.

Figure 1: Refugee and non-refugee immigrants by year



Notes: This figure shows the number of refugees and non-refugee household heads in our sample granted visas between 1986 and 2016. Non-refugee household head counts include all visa types.

Table 1: Descriptive Statistics

	Before 1999		After 1999	
	Refugees	Non-Refugees	Refugees	Non-Refugees
Female	49%	54%	39%	53%
Num. of Children	0.62	0.25	0.58	0.13
Family size	1.87	1.46	1.80	1.27
Married	47%	24%	52%	16%
Age	30.39	27.84	32.63	26.05
Count	27532	62091	6331	29965

Notes: This table shows descriptive statistics for immigrant household heads granted their first Danish visa from 1986-2003. Statistics refer to the averages in the first year in Denmark. Non-refugees include all visa types.

[13]A number of factors play a role in the staying rate of non-refugee immigrants versus that of refugees. In contrast to refugees, non-refugee immigrants have the option of returning to their home country. In addition, compared to refugees, non-refugee immigrants may benefit from more moving opportunities across commuting zones within Denmark. On the other hand, since a refugee's preferred residence location is likely different from the location he is assigned by the Danish Immigration Service, a refugee may have more incentive to ultimately leave than a non-refugee immigrant who chooses his first residence location himself. Overall the data indicate that the net effect of these forces is that refugees stay longer than non-refugee immigrants.

Table 2: Initial Location

	Before 1999		After 1999	
	Refugees	Non-Refugees	Refugees	Non-Refugees
Copenhagen	26%	53%	23%	52%
Bornholm	0%	0%	1%	0%
West and South Zealand	11%	5%	14%	4%
Fyn	11%	7%	9%	6%
Central and South Jutland	31%	23%	27%	25%
West Jutland	11%	5%	13%	5%
North Jutland	11%	6%	12%	7%
Count	27476	61977	6316	29880

Notes: This table shows the Danish geographical region where immigrant household heads first granted visas in the period 1986-2003 resided in their first year in Denmark. While our analysis is based on commuting zones, this table is based on somewhat larger Danish geographical regions to comply with disclosure rules. All visa types for non-refugee household heads are included.

Table 3: Visa Types Percentages

Type of visa	Before 1999	After 1999
Work	9%	7%
EU	48%	37%
Family reunification	7%	2%
Student	30%	50%
Other types	6%	4%
Count	11842	29965

Notes: This table shows the number of first visa types granted to non-refugee immigrant household heads granted visas in the period 1997-2003. We do not have information on visa type before 1997.

4 Empirical Strategy and Findings

As we discussed in Section 2, the assignment of refugees to locations was random only conditional on a number of factors such as country of origin, marital status, and the year of refugee status. In addition, refugees were asked in placement interviews about their educational and medical needs as well as location preferences. Before 1999, refugees assigned to unattractive areas were even able to refuse the assignment and be reassigned, although this rarely happened in practice.

We choose a baseline sample intended to make refugee allocations as close as possible to a random assignment. First, we include immigrants granted first visas from 1986 to 2003.[14] This allows us to look at outcomes up to thirteen years after the refugee visa was granted, that is, outcomes ten years after the longest years of stay as required by the policy change in 1999.

[14]There is no attrition in our main empirical exercise by construction. All controls are measured at the time the residence permit was granted, so we observe these for all immigrants. The main outcome variable is a dummy which takes the value of one if an immigrant is still in her placement location. We record a zero for this outcome variable if the immigrant either resides in another commuting zone in Denmark or if she drops out of the data for any reason. This is important to avoid selection, as non-refugee immigrants are more likely than refugees to leave Denmark rather than move to another Danish commuting zone.

Table 4: Unconditional staying probabilities

	Before 1999		After 1999		
	Refugees	Non-Refugees	Refugees	Non-Refugees	Diff-in-Diff
After 3 years	75.3%	38.5%	88.0%	32.8%	18.4%
After 8 years	60.0%	19.6%	60.0%	16.0%	3.7%
After 13 years	51.7%	15.4%	50.2%	11.6%	2.3%

Notes: This table shows the unconditional fraction of refugees and non-refugee household heads granted visas in the period 1986-2003 who remain in their first observed commuting zone after three, eight, and thirteen years. The last column is difference-in-differences for probabilities of staying after 1999 relative to before, between refugees and non-refugee immigrants. The observation number for each row is 125,918. All visa types for non-refugee immigrants are included.

In our baseline sample, we do not include immigrants with student or family visas because students have different profiles than refugees and people with family connections in Denmark are strongly tied to particular locations. We omit Copenhagen, because while many non-refugee immigrants choose to first live there, only refugees with special needs are placed there. Finally, we only include refugees placed in the last six months of the calendar year after the policy change in order to avoid those that were able to choose their settlement location when the quotas were not yet filled.[15]

To identify root growing—that staying in a location causes a person to stay there longer—we consider three different designs: regression discontinuity in Section 4.1, difference-in-differences in Section 4.2, and duration modeling in Section 4.3. We discuss how these three designs complement each other in Section 4.4, present a number of robustness checks in Appendix B, and discuss threats to identification in Appendix C. We provide suggestive evidence for the potential mechanisms behind our results in Section 5.

4.1 Regression Discontinuity

We consider a regression discontinuity design centered around the policy change on January 1st, 1999. The object we are trying to estimate in our regression discontinuity exercise is:

$$\beta_{rd} = \mathbb{E}\left[Y^1_{i,\tau} - Y^0_{i,\tau} | t^{visa}_i = t_0\right] \tag{1}$$

Here $Y^1_{i,\tau}$ is an indicator variable which equals one if individual i would have still been in the commuting zone of her placement municipality after τ years in Denmark if she were treated by the post-1999 policy. Similarly, $Y^0_{i,\tau}$ is an indicator variable which equals one if individual i would have still been in the commuting zone of her placement municipality after τ years in Denmark if she were *not* treated by the post-1999 policy. t^{visa}_i refers to the *day* on which the visa was granted, and t_0 is January 1st, 1999.[16] As is standard in the

[15] In appendix Section B we examine how the results are affected by cutting the data in different ways.

[16] In our other designs, the date on which visa was granted refers to the corresponding year rather than day.

quasi-experimental literature, the challenge is that we only observe people in either the treated state or the untreated state.

The identification assumption in this section is that there was no abrupt change in the characteristics of refugees who were granted visas at the same time as the policy change. If this assumption is valid, β_{rd} can be interpreted as the causal effect of the 1999 policy change on the probability that a refugee is still in her assigned commuting zone τ years after her visa was granted. We argue that the relevant aspect of the policy change for this outcome is that refugees were required to stay in their placement location for three years, rather than merely encouraged to stay for eighteen months.[17] If the refugees grow roots during the required three-year stay, we would expect them to be more likely than refugees who were not required to stay three years to remain in the commuting zone in the long run. In other words, our hypothesis is that β_{rd} is positive.

Under the identification strategy described above, we can estimate β_{rd} with the following estimator:

$$\begin{aligned}
\hat{\beta}^{+} &= \arg\min_{\beta_0^{+},\beta_1^{+}} \sum_{i=1}^{n} \mathbb{1}\left(t_i^{visa} >= t_0\right)\left(Y_i - \beta_0^{+} - \beta_1^{+}\left(t_i^{visa} - t_0\right)\right)^2 K_{bw}\left(t_i^{visa} - t_0\right) \\
\hat{\beta}^{-} &= \arg\min_{\beta_0^{-},\beta_1^{-}} \sum_{i=1}^{n} \mathbb{1}\left(t_i^{visa} < t_0\right)\left(Y_i - \beta_0^{-} - \beta_1^{-}\left(t_i^{visa} - t_0\right)\right)^2 K_{bw}\left(t_i^{visa} - t_0\right) \\
\hat{\beta}_{rd} &= \hat{\beta}_0^{+} - \hat{\beta}_0^{-} \qquad (2)
\end{aligned}$$

Intuitively, we estimate a weighted linear regression on distance from the time of policy implementation, with weights given by K_{bw}, which is a kernel function with bandwidth bw. We use a triangular kernel, and bandwidth chosen by the method suggested in Calonico et al. (2014), and report their suggested robust confidence intervals in our results.[18] As before, we run regressions for the dependent variable for tenure τ ranging from two to thirteen years.

Focusing on tenure $\tau = 13$, Figure 2 contains a bin plot showing how mean staying probabilities vary around the threshold, with each of the bins corresponding to twenty five days. Staying probabilities were falling up to the change in policy, and then jumped at the policy change. Recall that we are not controlling for the country of origin or first commuting zone in this specification, so it is possible that the mix of refugees and placement commuting zones was changing during this period towards people from countries of origin

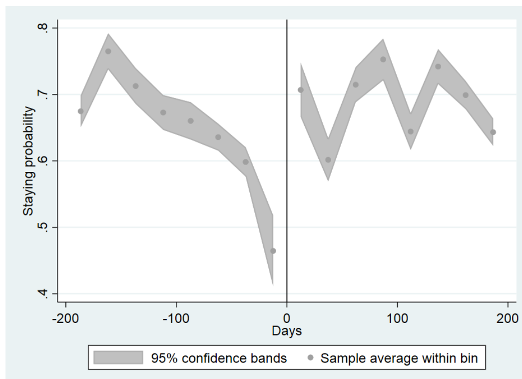
[17] We describe other aspects of the policy change in Section 2.2, and discuss implications of other aspects of the policy change for the interpretation of our results in Appendix Section C.

[18] We implement this method using the Stata package rdrobust on our entire sample. Obviously we cannot look only at refugees arriving in the final months of the calendar year as we did in our difference-and-difference exercise, since our regression discontinuity is calculated on the days surrounding January 1st. While the most recent version of rdrobust allows for the inclusion of controls, we found that including a large number of fixed effects is infeasible. Dropping those that were placed in Copenhagen and including only small dimensional controls has little effect on the estimates.

which were less likely to remain and less attractive placement commuting zones.

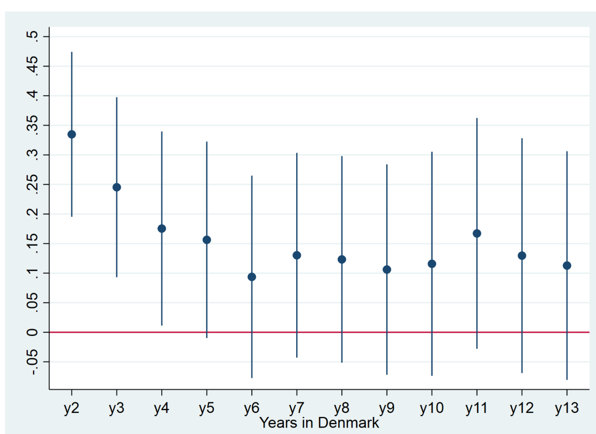
Our full estimates are presented in Figure 3, including 95% robust confidence intervals. The bandwidths calculated using the Calonico et al. (2014) method vary from year to year, ranging from 151 to 221 days around the cutoff. The point estimates indicate that those who arrived after the policy change were 11.3 percentage points more likely than those who arrived before the policy change to still be in their placement commuting zone after thirteen years in Denmark. The point estimates are noisy, and are either marginally or not significant at the 5% confidence level.

Figure 2: Regression Discontinuity, Binned Staying Probability 13 Years After Residence Granted



Notes: This figure shows mean staying probabilities for refugees granted residence in Denmark around January 1st, 1999. Each bin represents 25 days for 200 days before and after the policy change. The shaded area represents 95% confidence intervals for each bin mean. This plot is based on the 1,498 refugee household heads who were granted residence within 200 days of policy implementation.

Figure 3: Regression Discontinuity Estimates by Years After Residence Granted



Notes: This figure shows our regression discontinuity results from separate regressions in which the outcome of interest (i.e. whether a refugee is still in her initial location after τ years) ranges for τ between 2 and 13. Bars represent 95% confidence intervals clustered at the country of origin by first city level. These estimates are based on the 1550 refugee household heads who arrived within the estimated optimal bandwidth 211 days of the policy implementation. There are no controls in this specification.

4.2 difference-in-differences Regressions

The "difference-in-difference" design employs the following model:

$$Y_{i,\tau} = \beta_{dd} \mathbb{1}_{\{i=\text{refugee}\}} \times \mathbb{1}_{\{t_i^{visa}>=1999\}} + \gamma_r \mathbb{1}_{\{i=\text{refugee}\}} + \gamma_p \mathbb{1}_{\{t_i^{visa}>=1999\}} + X_i \delta + \epsilon_i \tag{3}$$

The dependent variable $Y_{i,\tau}$ is an indicator that equals one if person i was still in her first observed commuting zone τ years after her visa was granted in year t_i^{visa}.[19] $\mathbb{1}_{\{i=\text{refugee}\}}$ is an indicator that equals one if person i was a refugee, and zero otherwise (i.e. if i was a non-refugee immigrant). $\mathbb{1}_{\{t_i^{visa}>=1999\}}$ is an indicator that equals one if person i was granted a visa after the policy change in 1999, and zero otherwise (i.e. if the visa was issued before 1999). Controls include fixed effects for age at the granting of the residence permit, gender, marital/family status at the granting of residence permit, first commuting zone by year of immigration linear time trends, and first commuting zone by country of origin fixed effects.[20] The coefficient of interest is β_{dd}.

The identification assumption in this section is that in the absence of the policy shock, refugee and non-refugee immigrants would have had common time trends in terms of our dependent variable. Our rationale for using non-refugee immigrants as a control group is that they are, like refugees, also affected by changes in Danish immigration policy. During the period following the 1999 reform, Danish immigration policy was made more strict for both refugees and immigrants (Hvidtfeldt and Schultz-Nielsen, 2018). In 2002 the length of residency required for a permanent residence rose from three to seven years (Larsen et al., 2018).[21] A 2007 reform introduced work requirements for permanent residence as well as a higher standard of Danish language ability (Arendt et al., 2022).[22]

While the common trend assumption is impossible to directly check, we can examine refugee and non-refugee time trends in the period before our policy experiment to provide suggestive evidence. We do this graphically in Figure 4. If the common trend assumption holds, we expect there to be no divergence in the trends of refugees and non-refugee immigrants before the policy was implemented. We estimate equation (3) including controls as well as year dummies and the interaction of year dummies with the refugee dummy.

[19] The exact amount of time since the granting of the immigrant's visa depends on the calendar date that the visa was granted. Our first observation of an immigrant's location is on January 1st in the calendar year after the visa was granted. The fraction of immigrants in their first observed location after three years then means the fraction of immigrants who are still in the first observed location when we observe them on January 1st between three and four years after their visa was granted.

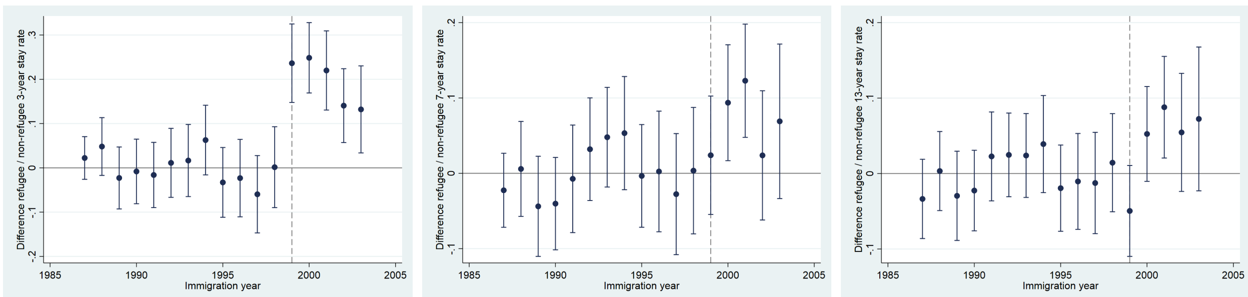
[20] We also use fixed effects for number of children and family size when we truncate the sample at 1997, because that is when this information becomes available in our data.

[21] Dustmann et al. (2022) find evidence that the residence permit type affects investments in host-country human capital.

[22] There are also challenges to the validity of the non-refugee group as a control group. For example, refugees are almost by definition more affected than non-refugee immigrants by push factors in the country of origin. On the other hand, both refugees and immigrants are affected by pull factors like a booming economy or a strong social welfare system.

In Figure 4, we plot the interactions, which can be interpreted as the conditional difference in the staying rates between refugees and non-refugee immigrants. We examine outcomes three, eight, and thirteen years after the granting of visas. That is, since the policy change required three years of forced stay, we add two increments of five years.[23] The dotted vertical line indicates 1999, the year of the policy change. The solid vertical lines are 95% confidence intervals.

Figure 4: Refugee to Non-refugee Difference in Staying in First Commuting Zone by Year of Visa



(a) After three years (b) After eight years (c) After thirteen years

Notes: This figure shows the difference in refugee and non-refugee rates of staying in their first location 3, 8, and 13 years after the granting of the residence permit, conditional on controls, and by year of immigration. Controls are fixed effects for age, gender, marital/family status (all upon the granting of the residence permit), first commuting zone by year of immigration linear time trends, and first commuting zone by country of origin fixed effects. Our sample for these figures omits non-refugee household heads with student or family reunification visas, anyone whose first observed location is Copenhagen, and refugees who were granted visas in the first six months of the calendar year after 1999. The total number of observations in each regression is 55,258.

The pre-policy trends for refugees and non-refugee immigrants appear to be similar, at least on average. After the policy change, there is a clear divergence in the expected direction. In particular, refugees who arrive after 1999 are more likely to still be in their first commuting zone in all plots. That the policy was immediately effective is suggested in subplot 4a. After the policy change, all cohorts of refugees are much more likely to remain in their placement commuting zone three years out relative to immigrants. In contrast, the relative outcomes of the two groups appear to be stable before the policy change. Outcomes after eight and thirteen years are similar, although the cohort granted visas in exactly 1999 at the onset of the policy change saw a decrease in their likelihood of remaining in their placement commuting zones relative to non-refugee immigrants.[24] Other cohorts of refugees react to the policy consistently with the hypothesis that refugees were more likely to remain in their placement city after the policy change.

We report in Table 5 the estimation results of our difference-in-differences equation (3) for $\tau = 13$ years after the visa was granted. The first two columns of the table contain results for our entire sample

[23]We cannot analyze later years without severely reducing our post policy sample size, since our final year of data is 2016. In any case, we will show results below which suggest that the estimated effect of the policy appears to be constant after around seven years in Denmark.

[24]Even without including any controls, members of the 1999 cohort are significantly less likely than preceding and following cohorts to remain in their placement commuting zone for 13 years. Experimenting with the data, we have found that including immigrants first observed in Copenhagen eliminates this unusual data point, but at the cost of creating other outliers.

period from 1986-2003, while the second two columns contain 1997-2003, the period for which we can more confidently identify both refugees and non-refugee immigrants. The coefficient on the refugee indicator, γ_r, is positive, sizable, and statistically significant in all specifications. For those who arrived before the policy change, refugees are 15.8-19.8 percentage points more likely to be in the placement commuting zone than non-refugees thirteen years after the visa was granted. On the one hand, we might have expected this coefficient to be negative, since non-refugee immigrants could select their preferred settlement location while refugees were initially conditionally randomly allocated. On the other hand, non-refugee immigrants had more moving options than refugees. In particular, they could move to a third country, or return to their home country. Since there are more moving options available to non-refugee immigrants, they are more likely to leave. It is an empirical question whether the former (selection) or the latter (moving options) is dominant. Here, we find that overall the latter is the dominant force.

The coefficient on Post-1999, γ_p, is negative across all specifications. The estimates imply that non-refugee immigrants were 1.5 to 3.2 percentage points less likely to remain in their first commuting zone for thirteen years if they arrived after 1999. To interpret this effect, recall that in this specification we omit immigrants whose first commuting zone was Copenhagen. The trend of leaving the first commuting zone may reflect the increasing attractiveness of the Copenhagen area relative to the rest of Denmark during the sample period as reflected by the growing share of overall Danish population living in Copenhagen.[25]

Our coefficient of interest, β_{dd}, the interaction between the refugee and post-policy indicators, is positive, statistically significant at a conventional five-percent significance level in three of our four specifications, and relatively large. We estimate this coefficient to be 4.8 percentage points at in our preferred specification under Column (2). We obtain similar estimates in the other specifications, except for the specification truncating the data from 1997 with controls where our estimate is lower at 2.19 percentage points and not statistically significant. In words, the 1999 policy change caused refugees to be two to five percentage points more likely to be observed in their settlement location thirteen years after receiving their visas. To put the magnitude of our difference-in-differences estimate in perspective, we note that the unconditional probability that a refugee is observed in his initial location after thirteen years is 51.7%. Based on our preferred specification, the policy caused a 9.3% increase of the baseline likelihood of staying thirteen years.[26]

[25] Our own calculations from Denmark Statistics data show that 41% of the Danish population lived in the Greater Copenhagen Area in 2008, and 44% in 2020.

[26] To relate our estimates to the literature that estimates moving costs, consider the following: Kennan and Walker (2011) report that a migration subsidy of \$10,000 would increase the fraction of households moving each year by two percentage points. Our preferred DiD estimates indicates that 4.8 percentage points more migrants stayed in their placement location after 13 years than would have if the policy had not been implemented. In our baseline sample, the average probability of staying for 13 years is 51.7%. Using the formula "probability of staying = $(1 -$ average probability of moving$)^{13}$", the average probability of moving in the baseline is 4.9%. In order to match the average probability of staying for those refugees affected by the policy change (56.5%), the probability of moving must be 4.3%. As a back-of-the-envelope calculation based on these numbers and those reported in Kennan and Walker (2011), the total effect of the policy was to create an implicit moving cost for refugees of

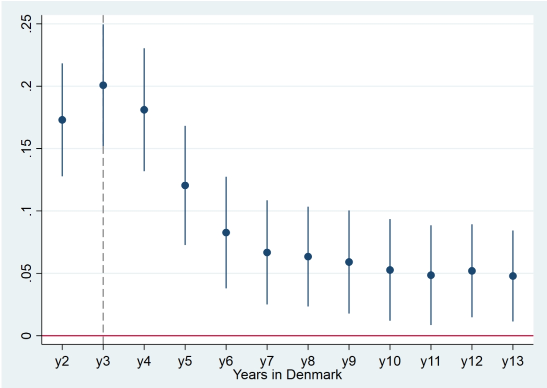
In addition, we run our difference-in-differences regression for each year at which we observe the outcome. In Figure 5, we plot the coefficient β_{dd} for analogues of Column (2) in Table 5, for τ ranging from 2 to 13 years. As can be seen in the figure, the maximum effect is three years after the visa is granted. This isn't surprising, since the policy essentially forces refugees to remain in their allocated commuting zone for three years. After the third year the effect falls in magnitude. The size of the effect is stable from year seven until year thirteen, suggesting that the long-run effect of the policy change was to cause refugees to be around five percentage points more likely to settle in their placement commuting zone.

Table 5: Estimation Results — difference-in-differences Specification

Dependent variable: Still in initial placement after thirteen years in Denmark				
	(1)	(2)	(3)	(4)
Refugee	0.163***	0.158***	0.171***	0.198***
	(0.0316)	(0.0282)	(0.0341)	(0.0404)
Post-1999	-0.0154	-0.0173	-0.0317***	-0.0244
	(0.0115)	(0.0134)	(0.0103)	(0.0153)
Refugee X Post-1999	0.0510***	0.0479**	0.0534***	0.0219
	(0.0191)	(0.0186)	(0.0168)	(0.0225))
Observations	55,258	55,258	12,339	12,339
R-squared	0.256	0.276	0.253	0.293
Truncate at 97	NO	NO	YES	YES
Controls	NO	YES	NO	YES
First CZ Trend	NO	YES	NO	YES

Notes: Standard errors clustered at the country of origin by first city level are in parentheses. Significance levels are indicated by *** as p<0.01, ** as p<0.05, * as p<0.1. All specifications include country of origin by first commuting zone fixed effects. Controls for full data include fixed effects for age, gender and family type (marital status) measured at the granting of the residence permit. If truncated at 1997, controls include family size and number of children measured at the granting of the residence permit as well. We omit immigrants whose first commuting zone is Copenhagen and those with student or family reunification visas.

Figure 5: difference-in-differences Estimated Coefficient by Years from Residence Permit

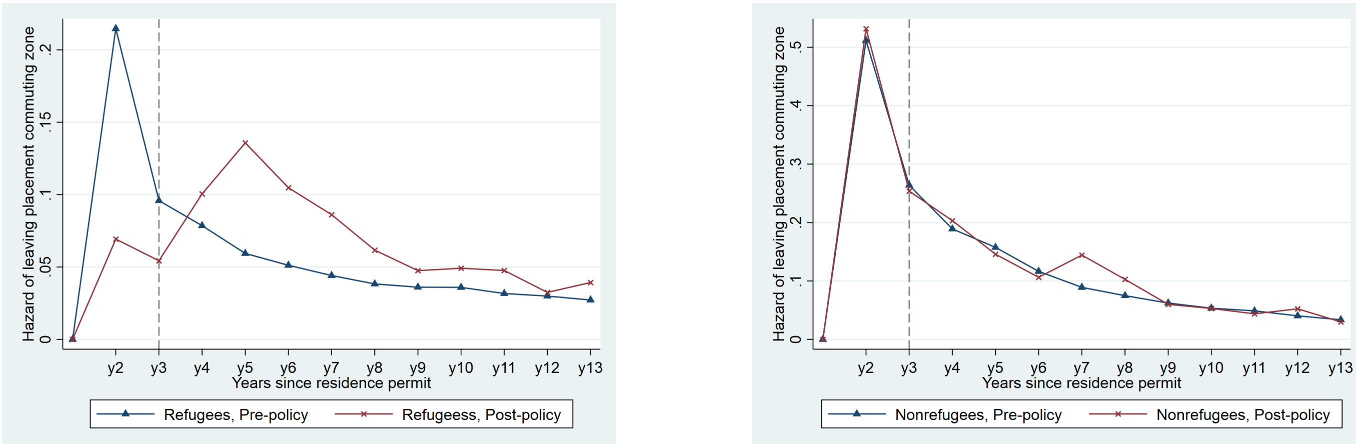


Notes: This figure shows the difference-in-differences coefficient, i.e. the interaction between refugee and policy indicators, from separate regressions for analogues of Column (2) in Table 5 with the outcome of interest (i.e. whether a person is still in her initial location after τ years) ranging for τ between 2 and 13. Bars represent 95% confidence intervals clustered at the country of origin by first city level. There are 55,258 observations in the underlying regression for each point in the figure.

$(4.9 - 4.3)/2 \times \$10,000 = \$3000$, or $1000 for each year the refugees were forced to stay in their placement location.

4.3 Duration Model

Figure 6: Duration Model—Hazard of Leaving First Commuting Zone



(a) Refugees (b) Non-refugee immigrants

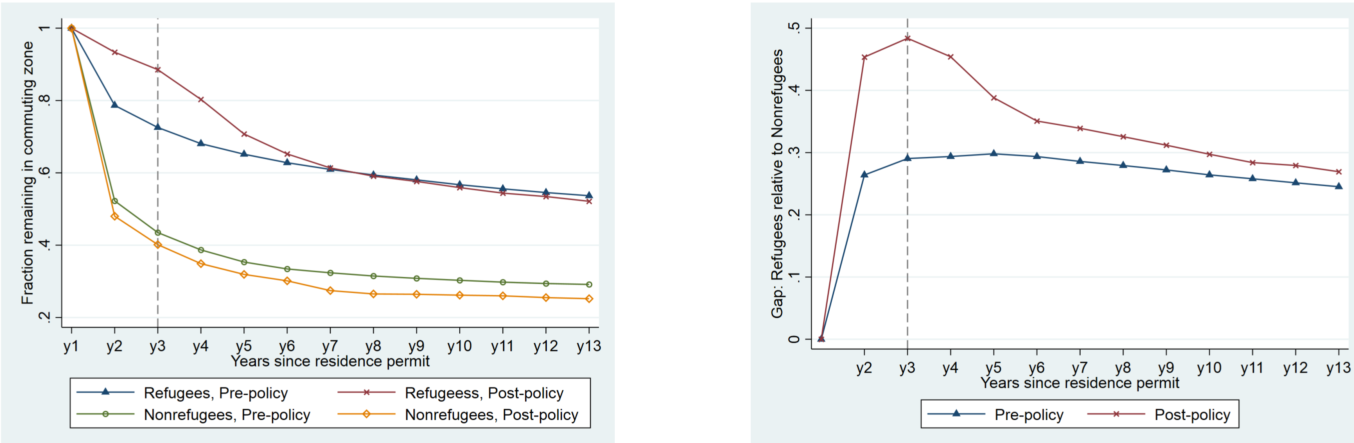
Notes: This figure shows the average estimated hazard rate of leaving the first observed commuting zone by duration year for a member of each of the following four groups: Refugees before and after the 1999 policy change in Panel (a), and non-refugees before and after the 1999 policy change in Panel (b). We omit immigrants whose first observed location is Copenhagen, student and family reunification visa holders, and refugees granted permits in the first half of the calendar year after 1999. There are 343,992 observations in our duration model regressions.

To complement our previous analyses, we use a duration model to further explore a person's decision to leave her location of placement. Specifically, we define the *survival rate* as how likely it is for a person to be still in her first commuting zone as a function of her duration of stay there. Correspondingly, the *hazard rate* for that person is defined as the likelihood that she leaves her first commuting zone conditional on a duration of stay there.

We use a discrete hazard model with a Probit link function. The duration model uses the same sample cut and controls for the origin country interacted with first commuting zone fixed effects as in the difference-in-differences specification above. The impact of the duration, τ, on the hazard of leaving is flexible, including indicator variables for each year of duration $\tau = 2, 3, ..., 13$. In other words, we do not impose proportional hazard rates across groups of immigrants, e.g., between refugees who arrived before and after the 1999 policy change.[27] We estimate the hazard rate for refugees and non-refugee immigrants who arrived before or after the policy change, conditional on the years they have stayed in their initial placement. Accordingly, our main independent variable is $\mathbb{1}_{\{i=\text{refugee}\}} \times \mathbb{1}_{\{t_i^{visa}>=1999\}} \times \mathbb{1}_{\{\text{duration}_i=\tau\}}$ where $\mathbb{1}_{\{i=\text{refugee}\}}$ is a dummy variable that equals one if person i is a refugee, $\mathbb{1}_{\{t_i^{visa}>=1999\}}$ is a dummy variable if person i arrived after the 1999 policy change, and $\mathbb{1}_{\{\text{duration}_i=\tau\}}$ is a dummy variable if person i's duration of stay equals τ years.

[27] To see why allowing for this flexibility is important recall from Figure 5 that in our difference-in-differences regression the coefficient of the interaction term (post-policy indicator interacted with the refugee indicator) is not monotonic in the first few years of duration. As we have extensively discussed, the reason is that the policy forced refugees to stay longer after the policy change, implying that the hazard rate is likely low for post-policy refugees in the first few years after receiving their residence.

Figure 7: Duration Model—Fraction Remaining in First Commuting Zone



(a) Survival rate

(b) Survival rate gap

Notes: In both sub-figures, pre-policy refers to the sample who were granted visas before 1999, and post-policy refers to that after 1999. Panel (a) shows the average estimated survival rate (as the fraction remaining in a commuting zone) by year of duration for each of the following four groups: Refugees before 1999, Non-refugees before 1999, refugees after 1999, and Non-refugees after 1999. Panel (b) shows the gap in the survival rate of Refugees relative to Non-refugees for the pre- and post-policy samples. We omit immigrants whose first observed location is Copenhagen, student and family reunification visa holders, and refugees granted permits in the first half of the calendar year after 1999. There are 343,992 observations in our duration model regressions.

Figure 6 shows the hazard rate for refugees (Panel a) and non-refugee immigrants (Panel b).[28] A few observations stand out. First, the hazard rate of leaving first commuting zone is notably larger for non-refugee immigrants. This observation confirms that non-refugee immigrants were more likely to leave their first commuting zone despite the fact that they chose the location themselves. Second, among non-refugee immigrants the hazard rates are similar between the group who arrived before and the group that arrived after the policy change. This observation, in turn, supports the assumption that non-refugee immigrants can be considered as a relevant comparison group in studying the impact of policy change on refugees' hazard of leaving.

Third, and most notably, the hazard rates are starkly different between refugees who arrived before the policy reform compared to refugees who arrived afterwards. Specifically, the hazard rate of pre-reform refugees monotonically falls with the duration of stay similar to the pattern for non-refugee immigrants, whereas the hazard rate of post-reform refugees is low in the first three years, peaks five years after they receive residence, and thereafter falls monotonically. The shape of the post-policy hazard function reflects how the reform made it difficult for refugees to leave their placement cities until after their third year there. The rise in moving hazard rates after the third year reflects a latent desire to move among refugees who were kept in their placement locations by the reform.

Using our estimated hazard rates, we then construct survival rates that inform about the cumulative

[28]Nielsen and Jensen (2006) contains a similar figure (their Figure 4.1) for refugees over a shorter time period.

impact of the 1999 policy change on location choices of immigrants. Panel (a) in Figure 7 presents the estimated survival rates, defined as the fraction of people who remained in their first commuting zone, for each of the four groups of refugees and non-refugee immigrants before and after the policy change. For non-refugee immigrants, the pre-reform group exhibits a higher survival rate, with the gap slowly widening as the duration of stay increased. Among refugees, the survival rate is initially larger for the post-reform group relative to pre-reform group, and as the duration of stay rises the gap shrinks until it virtually disappears after around seven years.[29] Put another way, comparing all refugees prior to the reform with all refugees after the reform, we see no difference in the long-run probability of remaining in the placement commuting zone.

However, under our difference-in-differences identification assumption, the gap in the survival rate of refugees before and after the policy change is not alone an indicative of root growing unless we compare it to an appropriate control group. This motivates Panel (b) of Figure 7, where we show the gap in the survival rate—the share remaining in first commuting zone—of refugees relative to non-refugee immigrants before and after the policy change. The figure shows that the gap narrows as the duration increases in the first few years, but even after 13 years, the gap exists. In other words, relative to the control group, a higher fraction of treated refugees tended to stay in their initial placement even after 13 years. Specifically, in the thirteenth year of duration, the gap equals 2.4 percentage points.[30]

4.4 Discussion

Each of the three empirical designs we estimate in this section share the same goal of identifying the causal effect of lengthening the mandatory stay in a placement municipality on the long-run probability of remaining in the commuting zone containing the municipality. Given the differences between our designs, it is not surprising that our results across these designs are not exactly the same.

The regression discontinuity compared to the difference-in-differences designs rely on different underlying control and treatment populations and on different identification assumptions. In particular, the regression discontinuity compares refugees who were granted visas just after the policy with refugees who were granted visas just before that. Since year-by-year macroeconomic fluctuations may affect refugees' staying rates, we restrict the sample to the refugees who were granted visas around ten months before and ten months

[29] By the nature of duration model estimation, once an immigrant leaves her placement location we no longer use her information in estimating hazards in subsequent years. With the same sample restrictions from our difference-in-differences estimation, we have relatively small samples especially for longer durations in the period following the reform. This can be seen in the somewhat bumpy hazard rate curve for refugees in the post reform period. We also estimate hazard rates on the unrestricted sample of immigrants and present the results in Figures A.5 and A.6 in the appendix.

[30] Estimating the duration model is computationally intensive, and the survival gap is a complicated function of cumulative hazard rates. While in principle standard errors could be bootstrapped, in this case it is not feasible due to the length of time required to estimate the duration model.

after the policy change. In comparison, in the difference-in-differences design, the identification relies on the assumption that, before the policy reform, staying rates of non-refugee immigrants were on a common trend with those of refugees from the same origin country and the same initial placement. Since the difference-in-differences design allows for a control group that is presumably affected similarly by year-to-year changes in economic conditions, we use a larger sample of those who were granted visas between 1986 and 2003.

There are advantages and disadvantages to each of our designs. The regression discontinuity design compares a refugee treatment group to a refugee control group, but it can only use information from a short window around the reform. Moreover, there is a worry that after the reform placements were less random, since refugees might be able to choose their placement location before quotas were filled. The difference-in-differences design and duration models allow us to use a longer time period, but require that refugees and non-refugee immigrants respond to macroeconomic shocks in a comparable way. This assumption is needed in order for us to control for the gradual tightening of immigration restrictions over that time period as well as other economic factors that could have similar impacts on refugees and non-refugee immigrants. Relying on the same sample, controls, and identifying assumptions as the difference-in-differences design, our duration model estimation produces hazard and survival rates around half as large as the difference-in-differences estimates. In terms of the design, the duration model is appealing because *(i)* it is fitting to think of the likelihood of leaving one's location as a hazard rate conditional on the duration of stay, which sets time as the state variable, *(ii)* it has a built-in capacity to take into account that errors are serially correlated, which is relevant since unobserved variables matter for who decides to stay in a location. On the other hand, because many immigrants have left their placement locations by the end of our sample, the hazard estimates for that period are likely to be noisy. In practice we find that the different designs, each relying on a different set of assumptions, deliver comparable results.

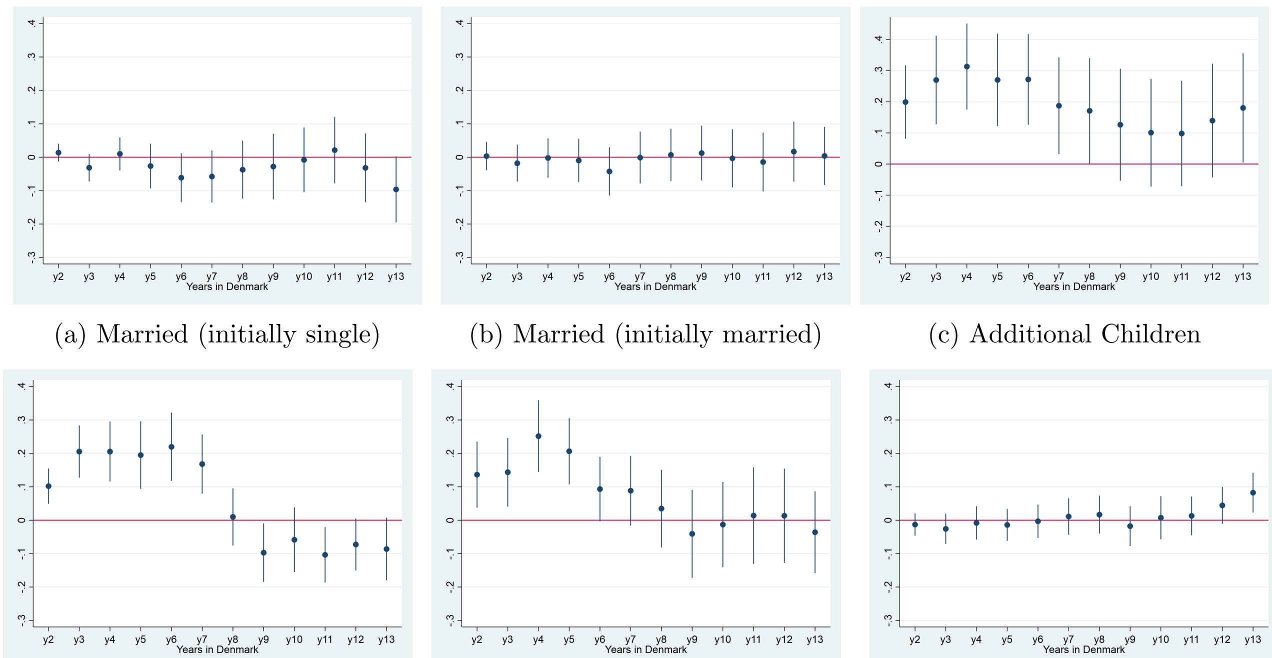
5 Mechanisms

Our main empirical results indicate a causal relationship from staying in a location on the likelihood of staying there longer, the phenomenon which we called root growing. What are the mechanisms through which staying in a location creates an attachment to that location, and how large are the opportunity costs of this attachment if one leaves her location? To study the mechanisms behind our results, we replace the dependent variable in our preferred difference-in-differences specification with a number of other individual characteristics. In a word, we are testing for the effect of being forced to remain in a location for three years on dependent variables such as employment or having more children.[31] If the policy had an effect on these

[31] An important caveat on these regressions is that we must drop observations for any immigrant who leaves Denmark, since we only observe individual characteristics such as employment as long as an immigrant remains in Denmark. This may introduce

variables, it may shed light on what mechanisms are behind the persistence in refugees' location choice. We consider the following outcomes: (a) whether a refugee is married if he was single when he received a residence permit, (b) whether a refugee is still married if he was married when he received a residence permit, (c) additional number of children relative to the first year of stay in Denmark, (d) whether a refugee is working if he was not working in his first observed year, (e) whether a refugee is working if he was working in his first observed year, (f) whether a refugee receives government support.[32]

Figure 8: Mechanism regressions



(a) Married (initially single) (b) Married (initially married) (c) Additional Children

(d) Working (initially unemp.) (e) Working (initially emp.) (f) Any government support

Notes: This figure shows the difference-in-differences coefficient, i.e. the interaction between refugee and policy indicators, with various outcome variables in separate regressions for years in Denmark ranging from one year to thirteen years. The outcome variables are: (a) whether the refugee is married for the sample who were single at the time he received a residence permit, (b) whether the refugee is still married for the sample who were married at the time they received their residence permits, (c) additional number of children relative to the first year in Denmark, (d) whether the refugee is working for the sample who were not working in the first year, (d) whether the refugee is working for the sample who were working in the first year, (f) government support. Except that we drop observations when they leave Denmark, the regression sample and specification is identical to that in Column 4 of Table 5 up to the outcome variables.

Figure 8 presents the results from this exercise in a form analogous to the way we presented the main results in Figure 5. In Panels (a) and (b), we see that the policy had no significant effect on marriage, neither for those who were single nor for those married at the time they received their residence permits. Panel (c)

selection into our estimates if the choice to leave Denmark is correlated with our outcomes of interest. We avoided this selection in our main exercises above by defining the outcome variable of interest as staying in the first observed commuting zone. If someone leaves Denmark, we know that they did not stay.

[32]We say a refugee is working if she has non-zero labor income. We say a refugee receives government support if he has non-zero income classified as either unemployment payments or welfare payments.

shows the effect of the policy on the number of children. Since we have fixed effects for the number of children at the time of residence permit, we are here examining additional children as the outcome. The estimates appear to be positive and statistically significant in the first few years, and positive and not statistically different from zero after a decade or so. Since the variable of children is a stock rather than a flow, this panel points to a change in the timing of fertility. Refugees forced to stay for three years may have had the same number of children as migrants who were not treated, but treated refugees had their children earlier.

Panel (d) and (e) show that treated refugees were more likely to find a job whether or not they were observed working in their first year. This effect is significant for the first few years, and fades away by the eighth year. We split the sample into those working in the first observed year and those not working in the first observed year because the control group of non-refugee immigrants may be more comparable if they have the same work status as refugees in their first year. The graph for the pooled sample looks similar. Panel (f) shows that the policy had little effect on government support take up.

Our results indicate that refugees who were forced to stay for three years were more likely to immediately have children and find work. It is plausible that there is a sunk cost to searching for work, and it makes sense to do so only if one plans on remaining in a location for a long enough period to cover the cost. Once one has a job, the job itself keeps one in a location. Moreover, coworkers naturally become friends.[33] Having children, by the same token, may create a natural social network of other parents for the refugees themselves, and the children may also develop friends and connections of their own as they get older. These social bonds may be costly to break by leaving the placement commuting zone. This interpretation is speculative, and we hope that future research can further illuminate the mechanism through which individuals grow roots.

6 Concluding Remarks

In this paper, we examine the hypothesis that people become more attached to a location the longer they live there. We are motivated to evaluate this root growing hypothesis because in the literature little is known about the mechanisms behind the persistence of cities and what causes path dependence in workers' locations, about forces behind moving costs, and why available estimates of moving costs are so large.

We show that refugees who were incentivized to stay in their placement municipality for three years after receiving their residence permits were more likely to be in the commuting zone of the municipality

[33]The policy change in 1999 which required refugees to remain in the settlement location for three years also extended integration coursework from one and a half to three years, and changed the bureaucracy surrounding refugee placement. We cannot rule out that these facets of the reform may have had a direct effect on job finding. For example, another possible reason that refugees are more often observed in the labor force after 1999 is that, as described in Section 2.2, refugees were allowed to move to a different municipality only if they had a job contract there. This may have given refugees more incentive to work in the formal sector rather than the informal sector, where they would not be recorded as working in our data.

thirteen years after their initial placement. This result holds in all three of the designs we examine—regression discontinuity, difference-in-difference, and a duration model. Our estimates are of varying sizes, and not all are statistically significant, but taken together they complement each other in supporting the root growing hypothesis. Additionally, we find suggestive evidence that treated refugees had children earlier and were more likely to be employed, which may have led them to form more links with the community. Earlier fertility and more intensive local job search are both choices, so they suggest investment rather than incidental effects of being forced to stay in the placement location. On the other hand, there is an interplay between incidental forces such as learning about one's current location and investment. If refugees foresee that incidental forces will keep them in the placement location, it will increase the returns to location-specific investment.

We hope that our work inspires future research in several directions. Related to the previous paragraph, researchers might further explore the mechanisms behind root growing. Is it occupational attachments, children's schools, or the network of friends that gradually tighten people's ties to locations? Additionally, developing and estimating a dynamic structural model of root growing might enable researchers to separate incidental root-growing from forward-looking location-specific investments. Finally, our results are based on the sample of refugees who arrived in Denmark between 1986 and 2003. It remains to be examined to what degree our results hold in other populations and time periods.

References

Allen, T. and Donaldson, D. (2022). The geography of path dependence. *Unpublished manuscript*.

Arendt, J. N., Bolvig, I., Foged, M., Hasager, L., and Peri, G. (2020). Language training and refugees' integration. Technical report, National Bureau of Economic Research.

Arendt, J. N., Dustmann, C., and Ku, H. (2022). Refugee migration and the labour market: lessons from 40 years of post-arrival policies in denmark. *Oxford Review of Economic Policy*, 38(3):531–556.

Autor, D. H., Dorn, D., and Hanson, G. H. (2015). Untangling trade and technology: Evidence from local labour markets. *The Economic Journal*, 125(584):621–646.

Azlor, L., Damm, A. P., and Schultz-Nielsen, M. L. (2020). Local labour demand and immigrant employment. *Labour Economics*, page 101808.

Baum-Snow, N. and Ferreira, F. (2015). Causal inference in urban and regional economics. In *Handbook of regional and urban economics*, volume 5, pages 3–68. Elsevier.

Bleakley, H. and Lin, J. (2012). Portage and path dependence. *The quarterly journal of economics*, 127(2):587–644.

Bosker, M., Brakman, S., Garretsen, H., and Schramm, M. (2007). Looking for multiple equilibria when geography matters: German city growth and the wwii shock. *Journal of Urban Economics*, 61(1):152–169.

Calonico, S., Cattaneo, M. D., and Titiunik, R. (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, 82(6):2295–2326.

Coate, P. and Mangum, K. (2019). Fast locations and slowing labor mobility.

Dalgaard, C.-J., Kaarsen, N., Olsson, O., and Selaya, P. (2018). Roman roads to prosperity: Persistence and non-persistence of public goods provision.

Damm, A. P. (2005). *The Danish Dispersal Policy on Refugee Immigrants 1986-1998: A Natural Experiment?* Department of Economics, Aarhus School of Business Aarhus.

Damm, A. P. (2009). Ethnic enclaves and immigrant labor market outcomes: Quasi-experimental evidence. *Journal of Labor Economics*, 27(2):281–314.

Damm, A. P. and Dustmann, C. (2014). Does growing up in a high crime neighborhood affect youth criminal behavior? *American Economic Review*, 104(6):1806–32.

Davis, D. R. and Weinstein, D. E. (2002). Bones, bombs, and break points: the geography of economic activity. *American Economic Review*, 92(5):1269–1289.

Dustmann, C., Adda, J., and Gorlach, S. (2022). The dynamics of return migration, human capital accumulation, and wage assimilation. *The Review of Economic Studies*.

Dustmann, C., Fasani, F., Frattini, T., Minale, L., and Schönberg, U. (2017). On the economics and politics of refugee migration. *Economic policy*, 32(91):497–550.

Dustmann, C., Vasiljeva, K., and Piil Damm, A. (2018). Refugee migration and electoral outcomes. *The Review of Economic Studies*.

Eckert, F., Hejlesen, M., and Walsh, C. (2022). The return to big-city experience: Evidence from refugees in denmark. *Journal of Urban Economics*, page 103454.

Edin, P.-A., Fredriksson, P., and Åslund, O. (2003). Ethnic enclaves and the economic success of immigrants—evidence from a natural experiment. *The quarterly journal of economics*, 118(1):329–357.

Foged, M. and Peri, G. (2016). Immigrants' effect on native workers: New analysis on longitudinal data. *American Economic Journal: Applied Economics*, 8(2):1–34.

Hvidtfeldt, C. and Schultz-Nielsen, M. L. (2018). refugees and asylum seekers in denmark. *Rockwell Fund Working Paper*.

Hybel, J., Jinkins, D., and Mulalic, I. (2023). Immigrants and the benefits of urban experience. working paper.

Kennan, J. and Walker, J. R. (2011). The effect of expected income on individual migration decisions. *Econometrica*, 79(1):211–251.

Kline, P. and Moretti, E. (2013). Local economic development, agglomeration economies, and the big push: 100 years of evidence from the tennessee valley authority. *The Quarterly Journal of Economics*, 129(1):275–331.

Krugman, P. (1991). History and industry location: the case of the manufacturing belt. *The American Economic Review*, 81(2):80–83.

Larsen, B., Kilstrom, M., and Olme, E. (2018). Should i stay or must i go?

Mian, A. and Sufi, A. (2014). What explains the 2007–2009 drop in employment? *Econometrica*, 82(6):2197–2223.

Michaels, G. and Rauch, F. (2018). Resetting the urban network: 117–2012. *The Economic Journal*, 128(608):378–412.

Miguel, E. and Roland, G. (2011). The long-run impact of bombing vietnam. *Journal of development Economics*, 96(1):1–15.

Ministry for Foreigners and Integration (2021). Boligplacering af nyankomne flygtninge. `https://web.archive.org/web/20210127042024/https://uim.dk/arbejdsomrader/Integration-af-nye-borgere/boligplacering-af-nyankomne-flygtninge`. Accessed: 2021-06-30.

Munch, J. R., Rosholm, M., and Svarer, M. (2006). Are homeowners really more unemployed? *The Economic Journal*, 116(514):991–1013.

Nakamura, E., Sigurdsson, J., and Steinsson, J. (2016). The gift of moving: Intergenerational consequences of a mobility shock. Technical report, National Bureau of Economic Research.

Nielsen, C. P. and Jensen, K. B. (2006). Integrationslovens betydning for flygtninges bosætning. *AKF report*.

Tombe, T. and Zhu, X. (2019). Trade, migration, and productivity: A quantitative analysis of china. *American Economic Review*, 109(5):1843–72.

Udlaendinge- og Integrationsministeriet (1998). Integrationslov, lov nr. 474.

Voigtlaender, N., Becker, S., Grosfeld, I., Grosjean, P., and Zhuravskaya, E. (2020). Forced migration and human capital: Evidence from post-wwii population transfers. *The American Economic Review*.

Yagan, D. (2014). Moving to opportunity? migratory insurance over the great recession. *Job Market Paper*.

Zabek, M. A. (2019). Local ties in spatial equilibrium.

A Data and Sample Selection

We primarily use two Danish registers. The first is the Residence Register. This register is created using information from the Danish Immigration Service on all first-time residence permits granted from 1997 until the present.[34] Important variables include the type of permit, the date the permit was granted, and a personal identification number of the recipient linkable across data sets. The second register we use is the Population Register. This register contains annual demographic information on all Danish residents beginning in 1986. Important variables for us are the municipality of residence, family/marital status, age, gender, and country of origin. We link the Population Register with the Residence Register using the personal identification number. In addition to these primary registers, we supplement our analysis with information from several other Danish registers, including information about family relationships from the Family Register and about income from the Income Register.[35]

Our sample contains only immigrants who were working age adults when the residence permit was granted, which we define as older than 18 and younger than 60.[36] Household head refers to a single adult not married to an existing Danish resident, not the child of an existing Danish resident, and not the parent of an existing Danish resident. In other words, we only analyze immigrants who had no close family connections in Denmark prior to the granting of their residence permit.[37]

An additional challenge with our data set is that we only observe first visa information for immigrants granted residence permits in or after 1997. Some of our empirical exercises use only this sample. In order to examine common trend assumptions and to get more statistical power, we sometimes use a larger sample including refugees arriving before 1997. In order to identify refugees granted permits in the period from 1986 to 1996, we follow a strategy used in much of the previous literature, and use immigrant origins to differentiate refugees from non-refugees (Damm, 2005; Damm and Dustmann, 2014; Dustmann et al., 2018).

[34]We obtain information on the first time someone is granted a residence permit within a permit type. For example, we observe the date at which someone is granted a family reunification visa, but not the date of renewal if that visa is renewed. If the same person later gets a work visa, we will observe the date that the work visa is granted.

[35]The Danish register acronyms are: Residence Register (OPHG), Population Register (BEF), Income Register (IND), and Family Register (FAM). Except for the exact date a visa was granted in the Residence Register, most other variables in our data set (e.g. location, marital status, number of children) are measured on January 1st. This means that we observe demographic characteristics of immigrants for the first time on January 1st in the calendar year following the year in which the visa was granted.

[36]Due to technical issues at DIS, data on immigration cases is not always linked to the ultimate personal identification numbers of residence permit recipients. In the Residence Register data, Denmark Statistics has made an effort to match missing personal identification numbers using nearest neighbor matching on attributes like country of origin, gender, age, etc. These methods work well when immigration flows are small and diverse, but not as well when immigrants are mostly homogeneous and from large source countries. In our analysis we drop these imputed observations, which represent 15.6% of the observations in the Residence Register. We also drop immigrants who we observe being granted a visa in a particular year, but who do not appear in our other registers until after the end of the following year.

[37]We do not observe more distant family relationships, such as whether an immigrant is the cousin or uncle of an existing Danish resident.

Each of these studies uses a slightly different set of origins. We use the list of refugee-sending countries from Damm (2009), which is based on the top 11 sources in the period 1986-1993, and add the former Yugoslavian republics since these were important sources in the mid-1990's. In practice, we classify any immigrant arriving from the following countries as a refugee before 1997: Lebanon, Iran, Iraq, Somalia, Sri Lanka, Vietnam, Poland (before 1990), Afghanistan, Ethiopia, Romania, Chile, Yugoslavia Federal Republic, Yugoslavia, Serbia & Montenegro, and Bosnia.

Our unit of geography is based on the commuting zones calculated in (Eckert et al., 2022). These commuting zones are collections of Danish municipalities. That paper, however, does not use data after 2007, when there was a merging of Danish municipalities into larger units. We modified the original commuting zones in two ways. First, 13 of the 273 old municipalities were split and assigned to different new municipalities. In this case we assign the entire old municipality to the new municipality with the lowest municipality code. The second issue is that some new municipalities contain old municipalities in different Eckert et al. (2022) commuting zones. In this case, we choose the modal commuting zone in each new municipality, and assign the old municipalities this commuting zone. This resulted in reassignment of 14 of the 273 old municipalities. Because of these two changes, our commuting zones will not be identical to the Eckert et al. (2022) commuting zones, but they will be consistent before and after the municipality reform.

In order to make all of our exercises comparable, we treat the first exit of an immigrant as a permanent exit, regardless of whether the immigrant later returns. This is required in our duration model, since it is estimated off of time to failure. This affects 2.2% of our observations. The regression discontinuity and difference-in-differences results are not sensitive to how we treat returns to the placement commuting zone.

B Robustness Checks

As with most observational studies, the number of specifications we could choose from is voluminous, both in terms of which controls we include, and exactly with which sample we run our analysis. As we argued in Section 2, refugees' assignment to locations was random only conditional on a number of control variables. In particular, across all of our specifications, we allow for country of origin by first commuting zone fixed effect. These fixed effects are important, because both before and after the policy change, refugees were more likely to be assigned to municipalities where a community from their country of origin already existed. That is, refugees might be assigned to commuting zones they prefer, invalidating the conditional random allocation assumption important for removing selection on location choice.[38]

[38]With our current fixed effects strategy, we exploit changes in behavior of all refugees relative to all non-refugee immigrants in response to the policy change, allowing for differences across mean staying rates for different countries of origin in different first commuting zones. We have experimented with including a three-way fixed effect of country of origin by first commuting

As a robustness analysis, we run our difference-in-differences regression for a wide range of alternative specifications. In Figure A.1, we plot the resulting coefficient, β_{dd} from regressions in which we set $\tau = 13$ and consider these alternative specifications:

- In regressions labeled as "Controls" we include age, gender, marital/family status, and if truncated at 1997, also the number of children, and family size fixed effects, all measured the first time they are observed, whereas in "No controls" we do not include them.
- Regressions labeled as "City trend" include a linear time trend in first commuting zone, and if "No City trend", then there is no trend included.
- In regressions labeled as "All mnths", the sample includes post-1999 refugees arriving in all calendar months, whereas in "Last 6 mnths" the sample includes only post-1999 refugees arriving in the last six months of the year. As discussed above, the distinction could matter since refugees might be able to influence their settlement locations only toward the first calendar months of a year when quotas were not yet filled.
- In regressions labeled as "All visas" the sample includes immigrants with all visa types, whereas in "No Stu/fam" we drop students and family reunification visas. We do so because students have very different moving pattern relative to other immigrants (such as a higher probability of moving to a new commuting zone after graduation), and family reunification visa receivers may be more tied to a location than refugees due to family already living there.
- In regressions labeled as "All cities" our sample includes refugees first observed in all commuting zones, whereas in "No Cph", we drop those whose we first observe in Copenhagen. As discussed earlier, Copenhagen was only available to refugees with special needs, for example special educational needs or certain medical conditions, none of which we observe in the data.
- In regressions labeled as "All prev yrs" the sample includes all years before 1997, while in "After 97 inc", we drop immigrants arriving before 1997. Identification of refugees is difficult before 1997 because we do not have direct information on visa type prior to 1997. We are more confident in our identification of refugees and non-refugee immigrants after 1997.

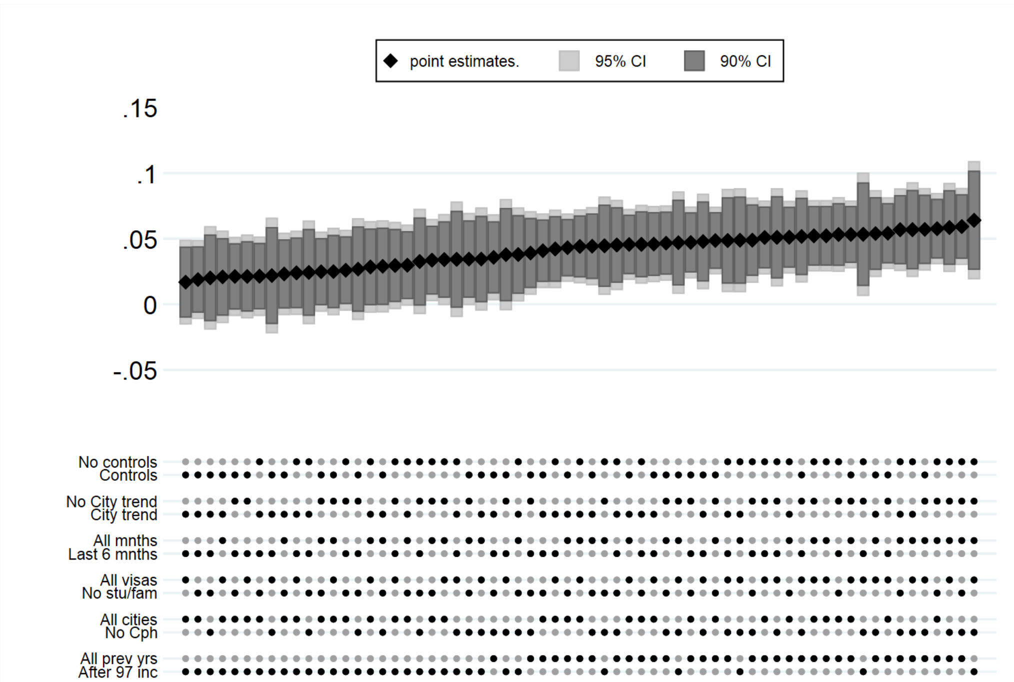
In each plot, we order the coefficients by point estimates. The coefficient of interest ranges from around two to seven percentage points, and is statistically significant in more than half of the specifications. There

zone by post-1999 dummy. With this three-way fixed effect, we would be identifying based on differences across mean staying rates of refugee and non-refugee immigrants from a country before and after the policy change. None of our results with that three-way fixed effect are statistically significant. The reason is simply that some countries (Afghanistan, say) overwhelmingly send refugees, and the other countries (Germany, for example) overwhelmingly send non-refugee immigrants. Hence, we have very little remaining variation within countries.

are not many obvious patterns among specifications which deliver low or high point estimates. The strongest pattern is that truncating the data to use only immigrants granted visas in 1997 and afterward tends to reduce the size of the coefficient.

We also include similar plots for outcomes of three and eight years after visas are granted (Figures A.2 and A.3). Three years after visas are granted we see a large policy effect (7 to 19 percentage points) across specifications. The results for eight years after visas are granted are roughly in line with those for thirteen years after visas are granted, with a bit higher variance in the point estimates.

Figure A.1: Diff-in-diff specifications, 13 years after visa was granted



Notes: This figure shows the difference-in-differences coefficient for a wide range of specifications with different controls and sample cuts. In front of each row below the graph, a black filled circle determines which feature is enabled in the corresponding specification.

C Threats to Identification

Our identification strategies rely on lengthening and strengthening of the requirement for refugees to stay in their placement location after the policy reform of 1999. As we described in Section 2, the change in policy involved a few other aspects alongside the longer required stay. While we acknowledge that the effects we find above are, in principle, the combined effects of all policy changes, we argue the most salient of these changes for the interpretation of our results is the lengthening of the required stay. Below, we specifically

Figure A.2: Diff-in-diff specifications, 3 years after visa was granted

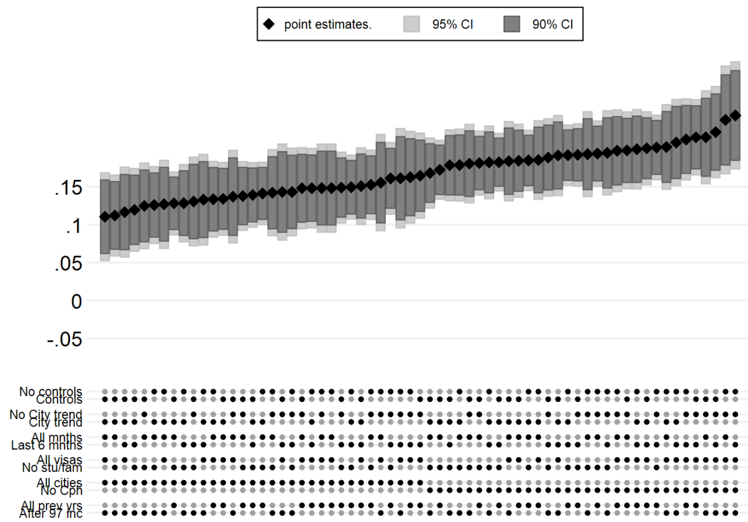
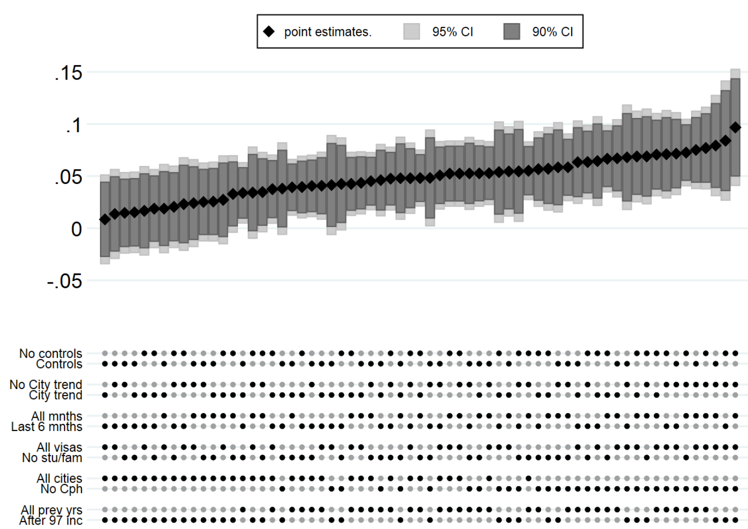


Figure A.3: Diff-in-diff specifications, 8 years after visa was granted



discuss three potential threats to our identification strategy.

First, a part of the 1999 reform was that the decision on where to place refugees passed from the non-profit DRC to the government office DIS. It is possible that DIS had different priorities in placing refugees, and may have placed them in different types of locations than DRC would have if the reform had not been undertaken. The primary goal of the policy was, however, the same both before and after the 1999 reform. According to Damm (2005), DRC aimed to distribute refugees across counties proportionally to population. The quota system introduced in 1999 pursued a similar goal. According to the website of the government section responsible for DIS, "the purpose of the quota system is to guarantee an even geographic distribution of foreigners in order to achieve better conditions for successful integration, as well as to guarantee that more municipalities participate in the task of integrating foreigners." (Ministry for Foreigners and Integration, 2021)[39]

Other aspects of the shift from DRC to DIS may have also biased our estimates. On the one hand, if refugees were placed in more desirable areas after the reform, then our estimates of root growing would be overestimates. For example, as we discussed earlier, refugees' location preferences might have entered into placement decisions after the policy reform. In the difference-in-difference and duration model sections of our analysis, we addressed this concern by restricting our sample to only refugee placements in the second half of each calendar year, when quotas for desirable locations were already filled. We could not use this strategy for the regression discontinuity design, because we must use placements on either side of the cut-off on January 1st.

On the other hand, several aspects of the reform might have led to less attractive placements for refugees, in which case our estimates would be underestimates. For example, as described earlier, the DRC placement policy before 1999 involved a rotating office which would place refugees arriving at the same time together. Since refugees at a given time were often fleeing from a particular event such as a civil war, this method of placement could lead to clustering of refugees by country of origin. To the best of our knowledge, placing refugees together by country of origin was not among the objectives of DIS after 1999. This potential discrepancy could have led to placements becoming less desirable for refugees after 1999.[40]

Second, it is a threat to our identification if unobserved characteristics of refugees or locations were different in systematic ways before and after the policy reform. Consider the case in which refugees who arrived after the reform differed systematically from those who arrived before the reform. Our controls, including

[39]This is the authors' translation from the Danish.

[40]In addition, some rural municipalities were not among the DRC placement locations because of the difficulty in finding housing there. These areas were not exempt from quotas after 1999. Presumably these rural areas were less desirable for refugees arriving in Denmark than other municipalities due to more limited work opportunities and offering a smaller set of product varieties.

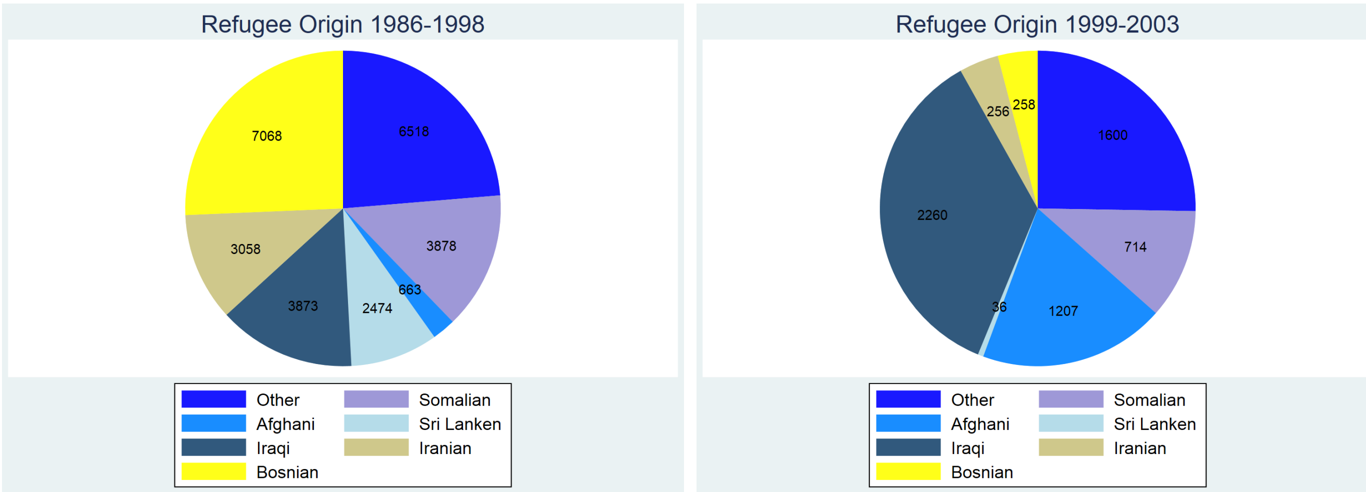
country of origin fixed effects, address this concern to a degree. Next, consider the case where local amenities evolved across locations at the same time as the policy reform we study. We include first-location linear time trends in our main specifications to partially address this concern. Additionally, we are encouraged by our regression discontinuity results, in which we have less power, but selection on unobservables is less likely.

A third threat to our identification is that the integration training offered to refugees was simultaneously lengthened from eighteen months to three years. For instance, Arendt et al. (2020) use a similar research design but focus on non-geographic outcomes. They find evidence that the additional language training in the 1999 reform caused refugees to have higher earnings. Was the reason refugees were more likely to remain in their placement location after 1999 their longer required stay, or was it their additional language coursework?

Since our identification is based on the timing of policy, and the lengthening of the required stay was simultaneous with the lengthening of the language training, it is difficult to separate these two aspects of the reform in the data. Nonetheless, Danish is spoken in all parts of Denmark. Additional language training should not have been a primary factor in a refugee's long-run decision to reside in one location rather than another within Denmark. In this regard, insofar as location is the outcome of interest, the lengthening of required stay is arguably the salient feature of the policy. We admit, however, that speaking better Danish might have led to more social and work opportunities, and those opportunities may have made it more likely for refugees to remain in their placement locations. While we cannot rule out this possibility, we hope that our paper motivates future research to separate these two channels.

D Appendix figures

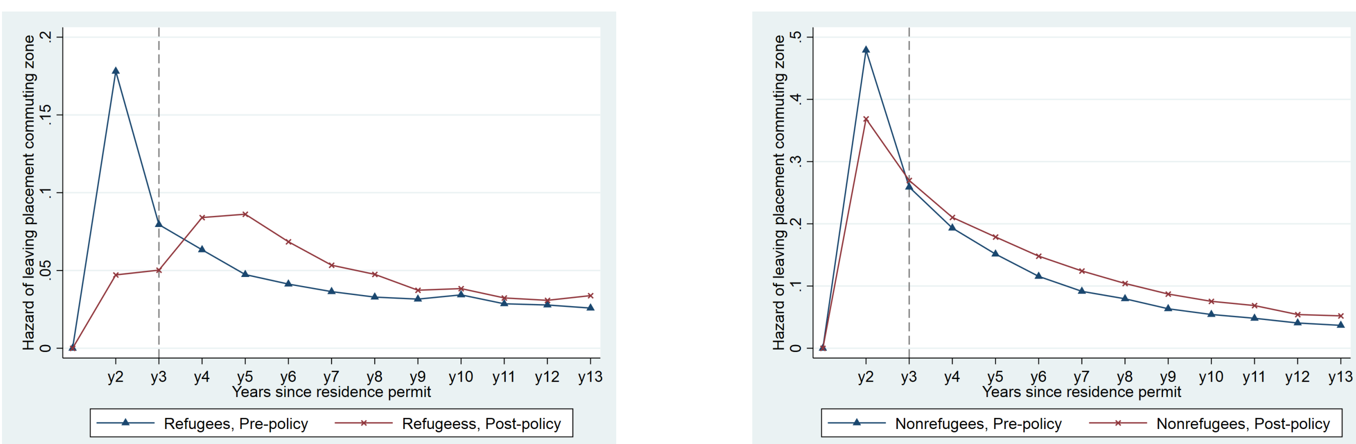
Figure A.4: Refugee countries of origin



(a) Visa granted pre-1999 (b) Visa granted post-1999 inclusive

Notes: This figure shows major countries of origin for refugee household heads who first received residence permits from 1986-2003. The total number of observations in the first panel is 27,532, and there are 6,331 observations in the second panel.

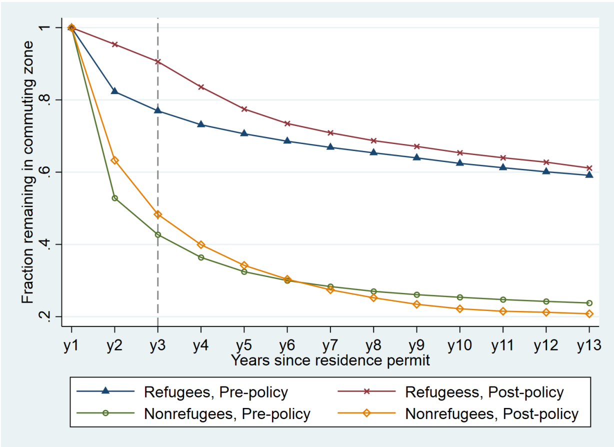
Figure A.5: Duration Model—Hazard of Leaving First Commuting Zone, Full sample



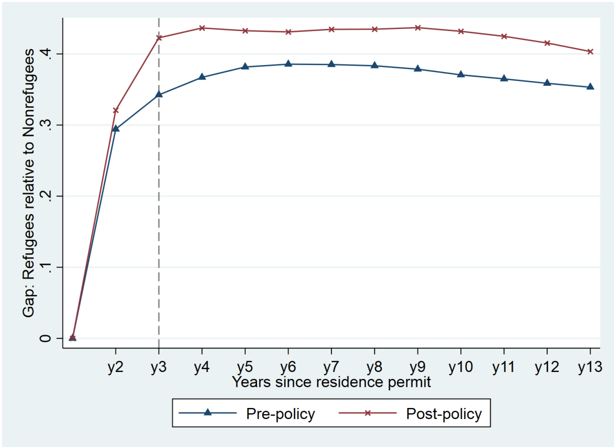
(a) Refugees (b) Non-refugee immigrants

Notes: This figure shows the average estimated hazard rate of leaving the first observed commuting zone by duration year for a member of each of the following four groups: Refugees before and after the 1999 policy change in Panel (a), and non-refugees before and after the 1999 policy change in Panel (b). We include our full sample of refugees and immigrants: all visa types, all first placement locations, and all calendar months of residence permits. There are 1,415,150 observations in our duration model regressions.

Figure A.6: Duration Model—Fraction Remaining in First Commuting Zone, Full sample



(a) Survival rate



(b) Survival rate gap

Notes: In both sub-figures, pre-policy refers to the sample who were granted visas before 1999, and post-policy refers to that after 1999. Panel (a) shows the average estimated survival rate (as the fraction remaining in a commuting zone) by year of duration for each of the following four groups: Refugees before 1999, Non-refugees before 1999, refugees after 1999, and Non-refugees after 1999. Panel (b) shows the gap in the survival rate of Refugees relative to Non-refugees for the pre- and post-policy samples. We include our full sample of refugees and immigrants: all visa types, all first placement locations, and all calendar months of residence permits. There are 1,415,150 observations in our duration model regressions.