MAX-PLANCK-INSTITUT FÜR SOZIALRECHT UND SOZIALPOLITIK
MAX PLANCK INSTITUTE FOR SOCIAL LAW AND SOCIAL POLICY

mea Munich Center for the Economics of Aging

The (Option-)Value of Overstaying

Romuald Méango

19-2020

MEA DISCUSSION PAPERS

mea – Amalienstr. 33_D-80799 Munich_Phone +49 89 38602-355_Fax +49 89 38602-390_**www.mpisoc.mpg.de**

The (Option-)Value of Overstaying

Romuald Méango

Abstract:

Asylum seekers with a rejected application account for three out of five illegal migrants in Germany. The objective of this paper is to calculate asylum seekers' ex ante returns on overstaying. It takes advantage of a unique survey designed to permit estimation of a structural model about the decision to overstay. The proposed model sees this decision as a utility maximization problem and translates it into a generalized Roy model with uncertainty about the final sector choice. As an important contribution to the literature, the paper proposes mild conditions for nonparametric identification of several objects of interest, including population distribution of ex ante pecuniary benefits, non-pecuniary costs, surpluses, and option-value created by the chance of becoming regularized in the future. Estimation is conducted using a semiparametric estimation procedure. Ex ante surpluses of overstaying are predominantly positive but very heterogeneous in the population. The pecuniary benefits explain a modest part of these surpluses. In contrast, the option-value is an important component of the expected returns. Afghan asylum seekers are ready to spend a long time with a precarious status to eventually obtain the right to stay and the amenities associated with it.

Zusammenfassung:

Asylsuchende mit einem abgelehnten Gesuch stellen drei von fünf illegalen Migranten in Deutschland. Ziel dieses Papiers ist es, die Ex-ante-Rendite von Asylsuchenden bei Überschreitung der Aufenthaltsdauer zu berechnen. Es stützt sich auf eine einzigartige Erhebung, die die Schätzung eines strukturellen Modells über die Entscheidung zum Überschreiten der Aufenthaltsdauer ermöglicht. Das vorgeschlagene Modell sieht diese Entscheidung als ein Problem der Nutzenmaximierung und übersetzt sie in ein generalisiertes Roy-Modell mit Unsicherheit über die endgültige Wahl des Sektors. Als einen wichtigen Beitrag zur Literatur schlägt das Papier moderate Bedingungen für die nichtparametrische Identifizierung mehrerer Objekte von Interesse vor, einschließlich der Verteilung von ex ante geldwerten Vorteilen in der Population, nichtmonetären Kosten, Überschüssen und dem Optionswert, der durch die Chance des zukünftigen Erhalts eines Aufenthaltstitels entsteht. Die Schätzung wird mit einem semiparametrischen Schätzverfahren durchgeführt. Die Ex-ante-Überschüsse von Aufenthaltsüberschreitungen sind überwiegend positiv, aber in der Bevölkerung sehr heterogen. Die geldwerten Vorteile erklären einen bescheidenen Teil dieser Überschüsse. Im Gegensatz dazu ist der Optionswert eine wichtige Komponente der erwarteten Erträge. Afghanische Asylsuchende sind bereit, lange Zeit mit einem prekären Status zu verbringen, um schließlich das Bleiberecht und die damit verbundenen Vergünstigungen zu erhalten.

Keywords:

Subjective expectations; Intention to overstay; Asylum seekers; Germany; Afghanistan

JEL Classification:

C20, D84, F22, J15, J18, J61, O15

The (Option-)Value of Overstaying *

Romuald Méango†

This version: December 3, 2020

Click here for the most recent version.

Abstract

Asylum seekers with a rejected application account for three out of five illegal migrants in Germany. The objective of this paper is to calculate asylum seekers' *ex ante* returns on overstaying. It takes advantage of a unique survey designed to permit estimation of a structural model about the decision to overstay. The proposed model sees this decision as a utility maximization problem and translates it into a generalized Roy model with uncertainty about the final sector choice. As an important contribution to the literature, the paper proposes mild conditions for nonparametric identification of several objects of interest, including population distribution of *ex ante* pecuniary benefits, non-pecuniary costs, surpluses, and option-value created by the chance of becoming regularized in the future. Estimation is conducted using a semiparametric estimation procedure. *Ex ante* surpluses of overstaying are predominantly positive but very heterogeneous in the population. The pecuniary benefits explain a modest part of these surpluses. In contrast, the option-value is an important component of the expected returns. Afghan asylum seekers are ready to spend a long time with a precarious status to eventually obtain the right to stay and the amenities associated with it.

Keywords: Subjective expectations; Intention to overstay; Asylum seekers; Germany; Afghanistan.

JEL codes: C20, D84, F22, J15, J18, J61, O15.

*Correspondence address: Munich Center for the Economics of Ageing. Amalienstr. 33, 80799 Munich, Germany. Tel: +49 89 38602 327. E-mail: meango@mea.mpisoc.mpg.de. This research was supported by the Max Planck Society for the advancement of science the research initiative "The Challenges of Migration and Integration", Grant M.FI.A.SOZ08002. I thank Axel Börsch-Supan, Mike Hurd, François Poinas, Andreas Steinmayr, and participants to various seminars for useful comments. The Institutional Review Board at the Max Planck Society for the advancement of science approved this study. All remaining errors are mine.

†Munich Center for the Economics of Ageing

1 Introduction

Between the years 2014 and 2016, the Federal Republic of Germany received nearly 1,100,000 registrations for asylum, the highest number in its history. As of 2019, more than 1,840,000 asylum seekers were accounted for in the country, among them 214,000 Afghan citizens, who form the second-largest group of asylum seekers after Syrians. Parallel to this surge, the number of foreigners with a rejected asylum application and a legal obligation to leave the country rose sharply, from 4,500 in 2014 to 152,000 in 2019. According to the population register, this accounts for three out of five irregular migrants. Afghan asylum seekers represent the largest subgroup in this category, with close to 25,000 individuals. Moreover, with 22% of asylum applications still in process and deportation to Afghanistan occurring seldom, this number is expected to rise over the next few years. While some have decided to remain, other Afghan citizens have already left the country. It is estimated that 5,580 left Germany in 2019, including 1,766 for whom an asylum status had been denied.

This research aims to shed some light on asylum seekers' perceived returns on overstaying. In particular, the objective is to estimate population distributions of *ex ante* benefits, costs, and surpluses associated with the overstay decision. Furthermore, it aims at understanding the importance of the option-value created by the chance of becoming regularized in the future. Finally, it seeks to measure the degree of uncertainty about returns in this population.

For this purpose, the research takes advantage of a unique survey designed to permit estimation of a structural model about asylum seekers' decision to overstay. The data collection was conducted during the second half of 2019 on Afghan asylum seekers in three large German cities: Berlin, Hamburg and Munich. More specifically, the survey elicited subjective beliefs about the chance of obtaining the right to stay in Germany (RtS) and the perceived risk of deportation. It also collected information on expected income and further outcomes depending on legal status. Finally, the survey elicited the intention to overstay under different hypothetical scenarios.

The proposed model sees the decision to overstay as a utility maximization decision as conceptualized by Sjaastad (1962). The maximization problem is then translated into a generalized Roy model with uncertainty about the final sector choice. In the model, the *ex ante* utility has three components: an expected pecuniary benefit, an expected non-pecuniary cost, and an idiosyncratic shock, also called "resolvable uncertainty." This uncertainty stems from the fact that beliefs are elicited prior to the actual decision. The *ex ante* utility of overstaying might be affected by additional shocks between the time of elicitation and the time of decision summarized in this parameter. Agents are then assumed to have probabilistic beliefs about the magnitude of this uncertainty. The stylized model allows deriving closed-form expressions for the objects of interest: distributions of *ex ante* benefits, costs, surpluses, uncertainty, and relevant elasticities.

As an important contribution to the literature, the paper proposes mild conditions for nonparametric identification of these objects of interest. Most studies using subjective expectations or choice probabilities to estimate structural models rely on parametric assumptions on the resolvable uncertainty parameter to achieve identification. Moreover, it is usually assumed that agents hold common beliefs about this uncertainty. This paper relaxes both assumptions; in particular, it allows for the uncertainty parameter to depend on individual-specific unobserved characteristics.

The study takes advantage of two features of the data to achieve nonparametric identification: a continuous dependent variable (intention to overstay, elicited as a choice probability) and repeated observation in counterfactual scenarios (with exogenous variation in the chance of becoming regularized). The proof of identification proceeds in two steps. In the first step, all objects of interest are expressed as functionals of a nonparametric reduced-form function of the intention to overstay. In the second step, the two aforementioned features of the data are exploited to propose a closed-form expression of the reduced-form function in terms of cumulative distribution functions (CDFs) of the observed data. Estimation is then conducted using a semiparametric procedure that alleviates the usual caveats associated with nonparametric estimation (the curse of dimensionality and large support requirement).

Intention to overstay is high in the population, with respondents reporting, on average, a 64% chance of overstaying without a RtS. This proportion varies considerably across cities of residence, with an average willingness to overstay that is 20 percentage points (pp) lower in Munich than in the two other cities. These intentions are highly sensitive to the perceived chance of becoming regularized in the future.

Given current beliefs about the chance of becoming regularized, *ex ante* surpluses of overstaying are predominantly positive but very heterogeneous in the population. The bottom 20% of the population expects negative surpluses on average. Pecuniary benefits explain a modest part of these surpluses, whereas the option-value created by the option to be regularized in the future is a key component of the *ex ante* returns. For the median individual, the option-value explains about 60% of the expected surplus. Furthermore, a one-percent decrease from the average probability of becoming regularized decreases, on average, by 0.39% the intention to overstay. This relatively strong elasticity emphasizes the importance of the chance of becoming regularized. Finally, the uncertainty parameter has quite a large variance, highlighting the high degree of uncertainty in the population.

Given the importance of the probability of being regularized, the paper studies its optimal choice by a social planner who wishes to minimize host-country's costs associated with overstaying migrants. The analysis shows that a sufficient condition for a large-scale regularization to be cost-efficient is that social costs associated with regularized migrants and pull-effect of regularization are small to moderate.[1] Moreover, when costs associated with regularized migrants are small, a large-scale regularization remains preferable to a restrictive policy even if the pull-effect is large. Finally, when a better chance of regularization increases asylum seekers' investments in country-specific human capital, a large-scale regularization becomes optimal for more combinations of social costs and pull-effect.

This paper relates to the literature using individual subjective expectations to estimate structural discrete choice models. Several investment decisions and behaviors have been investigated in this framework, including birth control choice (Delavande, 2008; Miller et al., 2020), risky sexual behavior (Delavande and Kohler, 2016), education choice (Jensen, 2010; Attanasio and Kaufmann, 2014), choice of college major (Wiswall and Zafar, 2015), and career decisions (Van der Klaauw, 2012). This research contributes to this growing literature by studying the subjective beliefs of asylum seekers into relation with their decision to overstay. Furthermore, it provides a novel and feasible identification and estimation strategy for several objects of interest, relaxing strong assumptions about the resolvable uncertainty parameter.

Within the literature on migration, the subjective expectation framework has been used to understand migrants' expectations about outcomes at destination (e.g., McKenzie et al., 2013; Hoxhaj, 2015). Closely related to the present paper, Mbaye (2014) and Bah and Batista (2018) look at the effect of individual expectations on the decision to migrate illegally. In particular, Bah and Batista (2018) emphasize the importance of the perceived chance of becoming regularized for the intention to migrate irregularly. Whereas previous contributions look at economic migrants, the focus of the current study is on a population of asylum seekers who have already arrived in the host country but face a significant risk of illegal stay. Asylum seekers with a rejected application form a large proportion of migrants with a legal obligation to leave Germany. The findings in this paper confirm the salience of subjective beliefs about the chance of becoming regularized in the future.

This paper is also related to the literature that studies the identification of the cost function in extended or generalized Roy models. This literature includes, among others, Bayer et al. (2011), d'Haultfoeuille and Maurel (2013), Eisenhauer et al. (2015), and Henry et al. (2020). These contributions entertain various covariate restrictions to achieve identification of the cost function. The present paper uses repeated observation of a continuous dependent variable under different hypothetical scenarios instead.

The two-step identification strategy developed in this paper is akin to the strategy proposed by Imbens and Newey (2009) for triangular simultaneous equations models without additivity. It avoids the need for parametric assumptions to estimate the whole decision model. Results on the identification of the reduced-form equation build on

[1] The pull-effect refers to the additional inflow of migrants that is generated by a generous regularization policy.

the contribution of Evdokimov (2010) that provides conditions for the identification of a nonparametric panel data model with unobserved heterogeneity. Bringing the identification results to the data, the paper builds on Chernozhukov et al. (2020) to propose a feasible semiparametric procedure that alleviates strong data requirements inherent to nonparametric estimation.

The rest of the paper is organized as follows. Section 2 gives a brief description of the context of asylum migration in Germany and the design of the survey. Section 3 presents a stylized model of the decision to overstay. Section 4 describes identification results. Section 5 discusses semiparametric estimation. Section 6 presents the empirical analysis. Section 7 discusses the social planner's problem. Finally, Section 8 concludes the paper.

2 Background

This section provides a brief contextual description of refugee migration to Germany and of the survey used in the empirical analysis. All sources for official statistics are collected in Appendix C.

2.1 Context of Refugee Migration in Germany

Between 2013 and 2019, the number of asylum seekers living in Germany tripled, from a little more than 615,000 in 2013, to more than 1,840,000 in 2019, with a peak of nearly 1,100,000 registrations between 2014 and 2016. As of 2019, Afghanistan, with 214,000 registered asylum seekers, was the second most important source country, before Irak (193,000) and after Syria (587,000).

Asylum seekers can file an initial application for asylum shortly after arrival. If rejected, they can file up to two subsequent follow up requests (appeal). The appeal process can last several months to several years. The outcome of an asylum application is uncertain for Afghan citizens. Between 2014 and 2019, one out of two applications was rejected in the first instance. As of 2019, two out of three applicants were recognized a protection status (92% of those as a temporary status with a maximum validity of three years), and 12% were denied and legally obliged to leave the country.[2]

Asylum seekers with a rejected application are asked to comply with a leave decree within a maximum period of 30 days and may receive financial support if they decide to leave voluntarily. If not complying, they face the risk of deportation. In practice though, deportation is rarely enforced. For example, in 2019, only 391 of the nearly 25,000 asylum seekers from Afghanistan with a legal obligation to leave Germany were returned to their home country, and 582 were returned to another European country under the

[2]In contrast, 96% of asylum seekers from Syria were recognized some protection status (with 97% as a temporary status), and only 1% were legally obliged to leave the country.

Dublin-agreement. Additionally, eight out of 10 Afghans who are legally obliged to leave Germany benefit from a temporary suspension of deportation or toleration status (in German, *vorübergehende Aussetzung der Abschiebung* or more simply, *Duldung*). This precarious status is issued when obstacles exist to deportation and can be valid for a period of a few days to a few months (usually not more than six months).[3] A *Duldung* does not provide the legal right to stay in Germany, has no guarantee of renewal and can be revoked under diverse circumstances.

Except under special circumstances, foreigners who have held a toleration status for at least three months can work in Germany if they receive a job offer and obtain approval of the Federal Employment Agency. According to the Asylum Seekers Benefits Act (*Asylbeweberleistunggesetz*), asylum seekers with a toleration status are entitled, during the first 15 months of their status, to receive social assistance to cover basic needs (food, accommodation, heating, health care, household consumption goods). After 15 months under a toleration status, the migrant is entitled to the same level of social assistance as a native.

Circumstances under which a toleration status can be transformed into a legal residence status include the completion of a qualified apprenticeship or study, or employment as a skilled worker for a two- to three-year uninterrupted period. Furthermore, in accordance with German law, if a foreigner cannot leave the country for a longer period of time for reasons beyond their control, he or she may be granted a residence permit for humanitarian reasons. However, this usually requires that the foreigner has a passport and has integrated into the local living conditions. This last condition is usually understood as showing proof of language proficiency and being able to provide for one's needs.

2.2 Survey Description

Within this context, the "Survey on Migrants' Expectations in Germany" was designed to understand the decision of Afghan asylum seekers to stay in Germany without the legal right to stay or exit to another country. Indeed, departure of Afghan citizens from Germany are not rare. An estimated 5,580 Afghans left Germany in 2019, including 1,766 cases where the asylum seeker had been denied protection. These numbers should be seen as lower bounds, as the exit is not always registered (e.g., when traveling by land). The survey elicited subjective expectations among Afghan migrants residing in the three German cities with the highest number of Afghan citizens: Berlin, Hamburg, Munich.

The elicited expectations can be divided into three categories: (i) subjective beliefs

[3]Opposing obstacles to deportations include, for example the right to safeguard the marital and family life or the assertion of illness-related dangers caused by deportation. A deportation is also impossible for factual reasons if travel documents are missing, the destination country refuses admission or traffic routes are interrupted. The immigration authorities also have the possibility of a discretionary tolerance for urgent humanitarian issues, personal reasons, or significant public interest (e.g. immediately upcoming surgery or the completion of a school or training year).

about population averages, (ii) subjective beliefs about individual outcomes if leaving or staying, and (iii) intention to overstay expressed as choice probabilities. Of particular interest in this study, the survey elicited expectations about the chance to obtain the RtS and the risk of deportation. More specifically, respondents were presented with the following hypothetical situations:
Imagine that your current status expires.

Q1. *What do you think is the percent chance that you would obtain the legal right to stay in Germany for the next three years?*

Q2. *You are not given the right to stay in Germany. But you decide to stay in Germany for the next three years. What do you think is the percent chance that you would obtain the legal right to stay in Germany by the end of the 3 years?*

Q3. *You live in Germany, but you do not have the legal right to stay in Germany. What do you think is the percent chance that you would be sent back to Afghanistan within the following three years?*

The 3+3-year window was selected for three reasons. First, it provides a time-horizon not too distant to form realistic expectations. Second, most protection statuses have a maximum validity of 3 years. Third, conversations with experts suggested that exit from a toleration status could be expected in a window of five to eight years.

The survey elicited further beliefs about outcomes in Germany and abroad, depending on the legal status of the individual. More specifically, respondents were asked about their average expected monthly income for the next three years with or without RtS.[4] Furthermore, the survey elicited beliefs about the perceived access to social services given their current status and in case of not obtaining a RtS.[5]

The main objective of the empirical analysis is to relate the above mentioned subjective beliefs to the intention to overstay. The survey elicited the intention to overstay as choice probabilities and in two formats. The first format was a direct question about the respondent's willingness to overstay:

Q4 *Imagine that your current status expired. You are not given the right to stay in Germany for the next three years. What do you think is the percent chance that you would decide to stay in Germany for the next three years?*

[4] The exact question was: "For each of the three situations, on average, what is the monthly income (including wage, government subsidies, etc.) that you expect you will have in the next 3 years (in euros)? Situation 1: Legal right to stay in Germany, Situation 2: without legal right to stay in Germany, Situation 3: You left Germany and live in another country because you do not have the right to stay in Germany."

[5] The perceived access was measured for four dimensions (education, labor market, social assistance and health services) on a Likert-scale: Full access, somewhat limited access, very limited access, no access at all.

Q5 *Suppose that you have been living for three years in Germany without the legal right to stay. If you are not given the right to stay after those three years, what is the percent chance that you will decide to stay in Germany for an additional three years?*

The second format presented the respondents with three hypothetical scenarios about the chance of becoming regularized:

Q6, Q7, Q8. *Imagine that your current status expired. You are not given the right to stay in Germany. But if you stay you will obtain with q% chance the right to stay in Germany at the end of the 3 years. What do you think is the percent chance that you would then decide to stay in Germany for the next 3 years?*

The parameter q was varied to take value 1, 50 and 99. All respondents received all three questions. The order of question was randomly assigned by the computer program.

The objective of the econometric and empirical analysis is to infer from the elicited subjective expectations *ex ante* costs and surpluses associated with the decision to overstay, and the degree of uncertainty in the population, as well as the elasticity of the overstay decision to subjective beliefs. To this end, section 3 introduces a stylized model that provides a closed-form expression of the objects of interest given the model's primitives.

3 Model

This section presents a simplified model of the decision to stay or leave Germany conditional on not having the legal right to stay (RtS). Appendix A discusses an extension. The general framework sees the decision to overstay as a utility maximization decision as conceptualized by Sjaastad (1962). Utility gains stem from differences in income, from private costs that can be money costs, or non-money costs (e.g., psychic costs), and from private returns.

The stay decision is described below as a generalized Roy model with uncertainty. The generalized Roy model is widely used in applied econometrics (see, for example, Amemiya, 1985; Heckman, 2001; Heckman and Vytlacil, 2005) and is the workhorse model for explaining migration decisions (see, for example, Borjas, 1987; Grogger and Hanson, 2011). The present model includes uncertainty about the final decision at the time of elicitation. The uncertainty stems from the fact that beliefs are elicited before the actual decision, and perceived utility might be affected by additional shocks between the time of elicitation and the time of decision. Agents are assumed to entertain probabilistic beliefs about the magnitude of these shocks.

Denote by i an asylum seeker currently living in Germany. Suppose two periods. In period 0, i does not receive the RtS in Germany. He can take one of two decisions: stay in Germany without the RtS or exit to another country. Exit is an absorbing state in that i will not return to Germany in the next period. If i decides to exit, he receives both in period 0 and in period 1 a monetary income and an additional utility that captures the amenities in the third country and the cost of moving. He forms beliefs about these quantities.

If i decides to stay, he or she faces a risk of deportation, associated with a cost of deportation. He or she has a chance of obtaining the RtS at the end of period 0 associated with a different stream of income and different amenities. The individual form subjective beliefs about these quantities. At the beginning of period 1, the individual is in either one of two states: with a RtS or without a RtS. If he obtains the RtS, he will almost surely stay in Germany. If not, he takes one of two decisions: stay in Germany without the RtS or exit to another country.

Denote by:

- Q_i: belief of individual i about the probability of obtaining the RtS at the end of period 0 after spending this period in Germany without the RtS (Q2).

- p_i^D: belief about the probability of being deported at the end of period 0 (Q3).

- P_{iQ}: intention to stay in Germany in period 0, conditional on not obtaining the RtS with Q percent chance at the beginning of period 0 (Q4, Q6-Q8).

- $P_i^{t=1}$: intention to stay in Germany in period 1 conditional on not obtaining the RtS at the end of period 0 (Q5).

The inter-temporal utility of staying in Germany in period 0 without RtS, $V_{i0}^{N,G}$, can be written:

$$V_{i0}^{N,G} = \alpha_i^{N,G} + \gamma_i \cdot Y_i^{N,G} + \nu_{i0}^G + \beta \mathbb{E} V_{i1}^{N,G} \tag{3.1}$$

where $\alpha_i^{N,G}$ represents i's expected (non-pecuniary) utility of residing in Germany without the RtS. $Y_i^{N,G}$ is the expected income without the RtS over period 0. γ_i represents the marginal utility of income and is assumed to be strictly positive ($\gamma_i > 0$). ν_{i0}^G is a utility "shock" that is unobserved by the agent at the time of the survey but will be observed at the time of decision. In the language of Blass et al. (2010), ν_{i0}^G is a "resolvable uncertainty", i.e., an uncertainty that will be resolved at the time the asylum seeker will actually decide to stay or leave. Finally, $\mathbb{E}V_{i1}^{N,G}$ is the continuation value, the expected utility in period 1.

A similar expression describes the inter-temporal utility of exiting at time 0:

$$V_{i0}^E = \alpha_i^E + \gamma_i \cdot Y_i^E + \nu_{i0}^E + \beta \mathbb{E} V_{i1}^E \tag{3.2}$$

where the terms are defined similarly. The expected utility in period 1 if staying in

Germany is described by the following:

$$
\begin{aligned}
\mathbb{E}V_{i1}^{N,G} \;=\;& p_i^D \cdot \left[\alpha_i^E + \gamma_i \cdot Y_i^E - c_i^D\right] \\
&+(1-p_i^D)\cdot Q_i \cdot (\alpha_i^{R,G} + \gamma_i \cdot Y_i^{R,G}) \\
&+(1-p_i^D)\cdot (1-Q_i)\cdot P_i^{t=1} \cdot (\alpha_i^{N,G} + \gamma_i \cdot Y_i^{N,G}) \\
&+(1-p_i^D)\cdot (1-Q_i)\cdot (1-P_i^{t=1}) \cdot (\alpha_i^E + \gamma_i \cdot Y_i^E)\,. \qquad (3.3)
\end{aligned}
$$

The first term on the right-hand side is the utility obtained if deported, an event that induces a cost c_i^D and is expected with probability p^D. The second term represents the utility when the agent receives the RtS at the end of period 1, which is expected with probability $(1-p_i^D)\cdot Q_i$. The third term is the utility when the agent does not receive the RtS but decides to stay in Germany, which is expected with probability $(1-p_i^D)\cdot Q_i \cdot P_i^{t=1}$. Finally, the last term is the utility attached to the decision of exiting in period 1, which is expected with probability $(1-p_i^D)\cdot Q_i \cdot (1-P_i^{t=1})$. Figure G.1 in Appendix G gives a tree-representation of the model.

The difference in utility between staying and leaving is given by the expression:

$$
V_{i0}^{N,G} - V_{i0}^E = \gamma_i \cdot B_i(Q_i) - \zeta_{i0} - \zeta_{i1}\cdot Q_i + \nu_{i0}^G - \nu_{i0}^E \qquad (3.4)
$$

where:

$$
\begin{aligned}
\zeta_{i0} \;=\;& \beta \cdot c_i^D \cdot p_i^D - (\alpha_i^{N,G} - \alpha_i^E)\cdot\left(1 + \beta\cdot(1-p_i^D)\cdot P_i^{t=1}\right) && (3.5)\\
\zeta_{i1} \;=\;& \left[P_i^{t=1}\cdot(\alpha_i^{N,G} - \alpha_i^E) - (\alpha_i^{R,G} - \alpha_i^E)\right]\cdot\beta\cdot(1-p_i^D) && (3.6)\\
B_i(Q_i) \;=\;& (Y_i^{R,G} - Y_i^E)\cdot\left[\beta\cdot(1-p_i^D)\cdot Q_i\right] \\
&+(Y_i^{N,G} - Y_i^E)\cdot\left(1 + \beta\cdot(1-p_i^D)\cdot(1-Q_i)\cdot P_i^{t=1}\right) && (3.7)
\end{aligned}
$$

At the beginning of period 0, ν_i^G and ν_i^E are observed by the individual, who chooses a location (stay in Germany or exit) to maximize the expected inter-temporal utility. Hence, i exits if and only if $V_{i0}^{N,G} - V_{i0}^{N,E} < 0$

Equation (3.4) can be rearranged to yield:

$$
\frac{1}{\gamma_i}\cdot\left(V_{i0}^{N,G} - V_{i0}^E\right) = \underbrace{B_i(Q_i)}_{\text{pecuniary benefit}} \; - \underbrace{\frac{1}{\gamma_i}\cdot\left(\zeta_{i0} + \zeta_{i1}\cdot Q_i + \nu_{i0}^E - \nu_{i0}^G\right)}_{\text{non-pecuniary cost}}, \qquad (3.8)
$$

which is the traditional representation of a generalized Roy model. In the terminology of this model (see for example Eisenhauer et al., 2015), $B_i(Q_i)$ is the *ex ante* pecuniary benefit; $C_{iQ} := \frac{1}{\gamma_i}\cdot\left(\zeta_{i0} + \zeta_{i1}\cdot Q_i + \nu_{i0}^E - \nu_{i0}^G\right)$ is the *ex ante* non-pecuniary cost when net utility shock is $\nu_{i0}^G - \nu_{i0}^E$. This cost can be positive or negative and summarizes variables other than expected income that influence the stay decision. Finally, $S_{iQ} := B_i(Q_i) - C_{iQ}$

represents the *ex ante* surplus when net utility shock is $\nu_{i0}^G - \nu_{i0}^E$. The individual stays in Germany if and only if $S_{iQ} \geq 0$.

Denote by ν_i the term $1/\gamma_i \cdot \left(\nu_{i0}^E - \nu_{i0}^G\right)$. C_{iQ} and, hence, S_{iQ} are unobserved by the agent at the time of elicitation. The agent is assumed to hold beliefs about the distribution of ν_i conditional on their information set. Denote by $F_{\nu_i|\mathcal{I}_i}$ this distribution, where $\mathcal{I}_i$ is the information set of individual i at the time of elicitation. At the time of elicitation, that is prior to the actual decision, the individual's expected probability to stay in Germany is then obtained as follows:

$$P_{iQ} = \Pr\left(S_{iQ} \geq 0 | \mathcal{I}_i\right) = \int 1\left\{S_{iQ} \geq 0\right\} F_{\nu_i|\mathcal{I}_i}\left(d\nu|\mathcal{I}_i\right) \tag{3.9}$$

ν_i represents the resolvable uncertainty in the model. There is no other resolvable uncertainty; the uncertainties about income, the probability of deportation, and the probability of obtaining the RtS are assumed to be unresolved by the time of location decision.

We are interested in the population distribution of subjective costs and surpluses and the distribution of uncertainty, and more generally, in the effect of policies that change beliefs about the probability of emigration. For each agent, one can define the individual's distributions of *ex ante* cost and *ex ante* surplus:

$$\begin{aligned} \Pr\left(C_{iQ} \leq c | \mathcal{I}_i\right) &= \int 1\{C_{iQ} \leq c\} F_{\nu_i|\mathcal{I}_i}\left(d\nu|\mathcal{I}_i\right) \text{ and} \\ \Pr\left(S_{iQ} \leq s | \mathcal{I}_i\right) &= \int 1\{S_{iQ} \leq s\} F_{\nu_i|\mathcal{I}_i}\left(d\nu|\mathcal{I}_i\right). \end{aligned}$$

or the individual's average *ex ante* cost and *ex ante* surplus :

$$\begin{aligned} \bar{C}_{iQ} &= \mathbb{E}\left(C_{iQ}|\mathcal{I}_i\right) = \int C_{iQ} F_{\nu_i|\mathcal{I}_i}\left(d\nu|\mathcal{I}_i\right) \text{ and} \\ \bar{S}_{iQ} &= \mathbb{E}\left(S_{iQ}|\mathcal{I}_i\right) = \int S_{iQ} F_{\nu_i|\mathcal{I}_i}\left(d\nu|\mathcal{I}_i\right). \end{aligned}$$

One may define further objects of interest such as derivatives of cost, surplus, and probability of emigration with respect to the individual's beliefs about regularization in the future, denoted by

$$\frac{\partial \bar{C}_{iQ}}{\partial q}, \frac{\partial \bar{S}_{iQ}}{\partial q}, \text{ and} \frac{\partial P_{iQ}}{\partial q}.$$

Finally, one can also define an individual-specific, *ex ante* option-value generated by the option of obtaining the RtS in the future:

$$\bar{S}_{iQ}|_{Q=q} - \bar{S}_{iQ}|_{Q=0} = \int_0^q \frac{\partial \bar{S}_{iQ}}{\partial q} dq$$

This parameter measures the increase in average *ex ante* surplus created by the chance of becoming regularized with probability q.

Table 3.1 provides closed-form expressions for the objects of interest according to

Objects of interest	closed-form expression when $\mathbb{E}\left(\nu_i \vert \mathcal{I}_i\right) = 0$
$\Pr\left(C_{iQ} \leq c \vert \mathcal{I}_i\right)$	$\Pr\left(\frac{1}{\gamma_i} \cdot (\zeta_{i0} + \zeta_{i1} \cdot Q_i) + \nu_i \leq c \vert B_i(Q_i), Q_i, \zeta_{i0}, \zeta_{i1}\right)$
$\Pr\left(S_{iQ} \leq s \vert \mathcal{I}_i\right)$	$\Pr\left(B_i(Q_i) - \frac{1}{\gamma_i} \cdot (\zeta_{i0} + \zeta_{i1} \cdot Q_i) - \nu_i \leq s \vert B_i(Q_i), Q_i, \zeta_{i0}, \zeta_{i1}\right)$
$\bar{C}_{iQ}$	$\dfrac{1}{\gamma_i} \cdot (\zeta_{i0} + \zeta_{i1} \cdot Q_i)$
$\bar{S}_{iQ}$	$B_i(Q_i) - \frac{1}{\gamma_i} \cdot (\zeta_{i0} + \zeta_{i1} \cdot Q_i)$
$\dfrac{\partial \bar{C}_{iQ}}{\partial q}$	$\dfrac{1}{\gamma_i}\zeta_{i1}$
$\dfrac{\partial \bar{S}_{iQ}}{\partial q}$	$\dfrac{\partial B_i(Q_i)}{\partial q} - \dfrac{1}{\gamma_i}\zeta_{i1}$
$\dfrac{\partial P_{iQ}}{\partial q}$	$\dfrac{\partial \bar{S}_{iQ}}{\partial q} \cdot F'_{\nu_i \vert B_i(Q_i), \zeta_{i0}, \zeta_{i1}, Q_i}\left(\bar{S}_{iQ} \vert B_i(Q_i), Q_i, \zeta_{i0}, \zeta_{i1}\right)$
$\bar{S}_{iQ}\vert_{Q=q} - \bar{S}_{iQ}\vert_{Q=0}$	$B_i(q) - B_i(0) - \dfrac{1}{\gamma_i}\zeta_{i1}\bar{q}$

Table 3.1: Parameters of interest expressed with the model's notation

the model notations when $\mathbb{E}(\nu_i|\mathcal{I}_i) = 0$. Note that each of these objects is defined at the individual level (hence the subscript i). The purpose of the empirical analysis is to learn about the distribution of these objects in the population, that is, if X_i is one of the parameters defined above, we wish to learn about its population distribution, $F_X(x)$.

4 Nonparametric Identification

A common approach to identification of structural models using subjective expectations (or choice probabilities) is to introduce parametric assumptions on the (joint) distribution of ζ_{i0}, ζ_{i1} and/or ν_i. In particular, most approaches assume common belief about the distribution F_ν in the population, usually assumed to be a normal or an extreme-value type 1 distribution (see, for example, Blass et al., 2010; Stinebrickner and Stinebrickner, 2014).[6] The approach to identification in this paper is nonparametric and imposes a milder condition on the distribution of the resolvable uncertainty. It takes advantage of two features of the data: a continuous dependent variable and repeated elicitation under different hypothetical scenarios.

The whole analysis, model, assumptions, and theoretical results in this section are understood to be conditional on a set of observed covariates (e.g., gender, age, city of residence) that are omitted from the notation hereafter unless they are involved in identifying assumptions.

The elicited intention to overstay can be written as:

$$P_{iQ} = \Pr\left(B_i(Q_i) - C(Q_i, \alpha_i, \nu_i) \geq 0 | B_i(Q_i), Q_i, \alpha_i\right) \tag{4.1}$$

Note that the probability measure is defined over ν_i. $C(Q_i, \alpha_i, \nu_i)$ represents the *ex ante* cost of staying without the RtS. This representation of the elicited decision can be viewed as an extension of the simple model used as motivation in Section 3. To see this, note that in this simple model, ν_i is additively separable, so that $C(Q_i, \alpha_i, \nu_i) = \bar{C}(Q_i, \alpha_i) + \nu_i$ and $\bar{C}(Q_i, \alpha_i) = \dfrac{1}{\gamma_i} \cdot (\zeta_{i0} + \zeta_{i1} \cdot Q_i)$. The transformation from a two-dimensional idiosyncratic term, (ζ_{i0}, ζ_{i1}), to a one-dimensional term, α_i, is always feasible, as noted by Chernozhukov and Hansen (2005). However, unlike (ζ_{i0}, ζ_{i1}), α_i does not have a structural interpretation. $\bar{C}(Q_i, \alpha_i)$ can be interpreted as a conditional quantile function and $\alpha_i \in [0, 1]$ as a conditional rank of individual i that summarizes the private information i holds about returns on overstaying. *A priori*, α_i is independent neither from $B_i(Q_i)$ nor from Q_i.

[6] Wiswall and Zafar (2015) use a linear fixed effect specification to net unobserved individual taste components that might shift the uncertainty parameter. Appendix F discusses the limitations of this procedure in our setup, and compares the results from the semiparametric estimation proposed below to the linear fixed effect estimation.

A reduced-form representation of Equation (4.1) is given by:

$$P_{iQ} = m(B_i(Q_i), Q_i, \alpha_i). \tag{4.2}$$

The proof of identification proceeds in two steps. First, it is shown that under a conditional independence restriction between ν and $(B(Q), Q)$ (Assumption 1 below), any object of interest in Table 3.1 can be expressed as a functional of $m(., q, a)$ at given values (q, a). Second, it is proved that the function $m(., q, a)$ and the (conditional) cumulative distribution function of α are identified under fairly mild restrictions.

4.1 Objects of interest as functionals of $m(., q, a)$ and F_α

Introduce the following restriction:

Assumption 1 (Independence of shocks)**.**

$$\nu_i \perp B_i(Q_i), Q_i | \alpha_i \tag{4.3}$$

This identifying assumption improves on traditional assumptions about the resolvable uncertainty in two ways. First, it does not assume a particular distribution for the purpose of identification. Second, unlike previous works that assume a common distribution of ν in the population, under Assumption 1, the distribution of beliefs remains individual-specific. Assumption 1 requires the existence of a scalar variable that captures all dependence between the resolvable uncertainty, expected income and the probability of becoming regularized.

Under Assumption 1, it will sometimes be convenient to rewrite S_{iQ} as follows:

$$\begin{aligned} S_{iQ} &= \bar{S}(B_i(Q_i), Q_i, \alpha_i) - V_{iQ}, \text{ where} \\ \bar{S}(b(q), \bar{q}, a) &:= b(q) - \bar{C}(q, a) \\ V_{iQ} &:= C(Q_i, \alpha_i, \nu_i) - \bar{C}(Q_i, \alpha_i) \\ \bar{C}(q, a) &:= \mathbb{E}_{\nu_i}\left(C(q, \alpha_i, \nu_i) | \alpha_i = a\right) \end{aligned} \tag{4.4}$$

When ν_i is additively separable, $V_{iQ} = \nu_i$ and $P_{iQ} = F_{\nu_i|Q_i,\alpha_i}\left(\bar{S}(B_i(Q_i), Q_i, \alpha_i)|\alpha_i\right)$. The key assumption to retrieve the above form is the following:

Assumption 2 (Strict Monotonicity)**.** *$C(q, a, \nu)$ is continuous and strictly monotonic in ν for all (q, a).*

Under assumptions 1 and 2, Equation (4.1) can be rewritten as:

$$P_{iQ} = F_{V_{iQ}|Q_i,\alpha_i}\left(\bar{S}(B_i(Q_i), Q_i, \alpha_i)|Q_i, \alpha_i\right) \tag{4.5}$$

Take for example $\Pr\left(C(q,\alpha,\nu) \leq c|\mathcal{I}_i\right)$, the individual-specific distribution of the *ex ante* cost when the chance of becoming regularized is q. Consider an individual with private information $\alpha = a$, who assumes a distribution $F_C(c;q,a)$ of the *ex ante* cost. Under Assumption 1, and for any c on the real line, $F_C(c;q,a)$ can be derived as follows:

$$\begin{aligned} F_C(c;q,a) &:= \Pr\left(C(q,\alpha,\nu) \leq c|\alpha = a\right) \\ &= \int I\left\{C(q,a,\nu) \leq c\right\} F_{\nu|\alpha}(d\nu|a) \\ &= \int I\left\{C(Q,a,\nu) \leq B(Q)\right\} F_{\nu|B(Q),Q,\alpha}(d\nu|c,q,a) \\ &= \int I\left\{B(Q) - C(Q,a,\nu) \geq 0\right\} F_{\nu|B(Q),Q,\alpha}(d\nu|c,q,a) \\ &= m(c,q,a) \end{aligned} \tag{4.6}$$

The third line makes use of assumption 1, which implies that $F_{\nu|\alpha}(d\nu|a) = F_{\nu|B(Q),Q,\alpha}(d\nu|c,q,a)$, for all c on the real line. The last line follows from the model (4.2). Equation (4.6) implies that the distribution of cost, $F_C(c;q,a)$, is identified on the conditional support of α given $(B_i(Q_i),Q_i)$. Our interest lies in its CDF, which is given by the following expression:

$$\int 1\left\{m(c,q,a) \leq l\right\} F_\alpha(da). \tag{4.7}$$

Identification of this CDF requires the following condition:

Assumption 3 (Support)**.** *For all $(B_i(Q_i),Q_i) \in \mathcal{B} \times [0,1]$, the support of α conditional on $(B_i(Q_i),Q_i)$ is the same as the support of α.*

Assumption 3 is similar to the support condition imposed by Imbens and Newey (2009) on the control function. It requires that variations of α that will be estimated from the observed data cover the whole support of this variable. If this condition fails, one can recover bounds on the CDF similar to Theorem 4 of Imbens and Newey (2009). The semiparametric model introduced for the empirical application satisfies Assumption 3.

More generally, several objects of interest take the generic form:

$$\int 1\left\{\Lambda(m(.,q,a)) \leq l\right\} F_\alpha(da) \tag{4.8}$$

where $\Lambda(m(.,q,a))$ is an integrable functional. Table 4.1 is parallel to Table 3.1 and summarizes all objects of interest as functionals $\Lambda(m(.,q,a))$.

All of the results are proved in the Appendix B. Note that under Assumption 2 and given knowledge of the average surplus function, the density of the resolvable uncertainty is identified from the following expression.

$$F'_{V|Q,\alpha}\left(\bar{S}(B(Q),Q,\alpha)|Q,\alpha\right) = \nabla_1 m(B(Q),Q,\alpha) \tag{4.9}$$

Parameter	$\Lambda(m(.,q,a)$
$\Pr\left(C(q,\alpha,\nu)\leq c\vert\alpha=a\right)$	$m(c,q,a)$
$\Pr\left(S(b(q),q,\alpha,\nu)\leq s\vert\alpha=a,b(q)\right)$	$1-m(b(q)-s,q,a)$
$\bar{C}(q,a)$	$\int_{\mathcal{C}^+}\left[1-m(c,q,a)\right]dc-\int_{\mathcal{C}^-}\left[m(c,q,a)\right]dc$
$\bar{S}(b(q),q,a)$	$b(q)-\bar{C}(q,a)$
$\dfrac{\partial\bar{C}(Q,\alpha)}{\partial q}$	$\dfrac{\nabla_q m(B(Q),Q,\alpha)}{\nabla_1 m(B(Q),Q,\alpha)}-\dfrac{\partial B(Q)}{\partial q}$
$\dfrac{\partial\bar{S}(B(Q),Q,\alpha)}{\partial q}$	$\dfrac{\nabla_q m(B(Q),Q,\alpha)}{\nabla_1 m(B(Q),Q,\alpha)}$
$\dfrac{\partial P_Q}{\partial q}$	$\nabla_q m(B(Q),Q,\alpha)$
$\bar{S}(b(q),q,\alpha)-\bar{S}(b(0),0,\alpha)$	$\int_0^q\dfrac{\nabla_q m(b(q),q,\alpha)}{\nabla_1 m(b(q),q,\alpha)}dq$

Table 4.1: Objects of interest expressed as functional of the reduced-form Equation (4.5)

In Section 4.3 below, it is shown that $m(.,q,a)$ and F_α are in turn identified from the data, so that the objects of interest are also identified from the observed data.

4.2 Alternative Derivation of Average Costs

Before turning to identification results for $m(.,q,a)$ and F_α, this section introduces an alternative condition to the support restriction (Assumption 3). As noted by Imbens and Newey (2009) and Florens et al. (2008), support restrictions can be very restrictive.

Assumption 4 (Beliefs normalization)**.** *There exists some $p_0\in(0,1)$, such that:*

$$F_{V|Q,\alpha}(0|q,a)=p_0,\ \text{for all } (q,a).$$

Assumption 4 is a location normalization. For example, when $p_0=0.5$, it implies that an individual that foresees on average neither loss nor gain from staying, i.e., $\bar{S}_{iQ}=0$, will report an equal chance to stay or to leave. Since the mean of V is also 0, it requires that mean and median are equal. Assumption 4 is satisfied for a zero-mean symmetric distribution. Symmetry of the resolvable uncertainty is an implicit assumption of most work that assume an extreme-value type I or normal distribution of the resolvable uncertainty. See also Blass et al. (2010) on the use of symmetry assumptions to achieve identification.

By assumptions 2 and 4, $p_0 = m\left(\bar{C}(q,a),q,a\right)$ for all (q,a). Provided that the function $m(y,q,a)$ is invertible in its first argument, it follows that:

$$\bar{C}(q,a) \quad = \quad m^{-1}(p_0,q,a) \tag{4.10}$$

Under Assumption 1 and Assumption 2, Equation (4.5) holds. Invertibility of $m(y,q,a)$ follows from the fact that the pecuniary surplus is separable from the cost.

4.3 Identification of $m(.,q,a)$ and $F_\alpha(.)$

The following identification results are closely related to Evdokimov (2010) that proposes identification results for a nonparametric panel data model with unobserved heterogeneity. In particular, the proof of identification follows Step 2 and Step 3 of Theorem 2 therein.[7] For each individual i, the following is observed:

$$P_{iq_i} \quad = \quad m(B_i(q_i),q_i,\alpha_i), \text{ for some } q_i \in \{0, 0.01, 0.02, \dots, 1\} \text{ and} \tag{4.11}$$
$$P_{iq_{0j}} \quad = \quad m(B_i(q_{0j}),q_{0j},\alpha_i), \text{ for } q_{01} = 0.01, q_{02} = 0.50, \text{ and } q_{03} = 0.99. \tag{4.12}$$

Let $\mathcal{Q}_i := \{q_i, q_{01}, q_{02}, q_{03}\}$. Denote by X_{iQ} the random vectors $(B_i(Q), Q)$, for $Q \in \mathcal{Q}_i$.

Assumption 5 (Restrictions on α and $m(x,\alpha)$)**.**
α and $m(x,\alpha)$ satisfy the following conditions:

(i) α_i is continuously distributed conditional on X_{iQ}.

(ii) For each x, there exists $\underline{\alpha}(x)$ and $\overline{\alpha}(x)$, $\underline{\alpha}(x) \le \overline{\alpha}(x)$, such that:

a. $m(x,\alpha)$ is strictly increasing in α, on $(\underline{\alpha}(x), \overline{\alpha}(x))$,

b. $m(x,\underline{\alpha}(x)) = 0$ and $m(x,\overline{\alpha}(x)) = 1$,

c. for all $\alpha < \underline{\alpha}(x)$, $m(x,\alpha) = 0$,

d. for all $\alpha > \overline{\alpha}(x)$, $m(x,\alpha) = 1$.

(iii) there exists $\bar{x} = (\bar{y},\bar{q})$, such that for all α, $m(\bar{x},\alpha) = \alpha$.

(iv) $m(x,\alpha)$, $f_{\alpha_i|X_{iQ}}(a|x)$, and $f_{\alpha_i|X_{iQ},X_{i\bar{q}}}(a|x_1,x_2)$ are everywhere continuous with respect to x, x_1, x_2, for all a, for $Q \in \{q_i, q_{01}, q_{02}, q_{03}\}$.

[7] Equation (4.2) in this paper differs from Equation (1) in Evdokimov (2010) in two ways: (i) there is no scalar idiosyncratic disturbance U_{it}, (ii) which allows including the "time" variable (in our case Q_i) to enter X_{it}. Appendix D deals with the case where there exists an additional idiosyncratic disturbance that is interpreted as measurement error.

Assumption 5.(i) is not restrictive since the function $m(x, \alpha)$ can be a step function in α. Monotonicity Assumption 5.$(ii).a.$ is standard in the analysis of nonparametric models and guarantees invertibility of function $m(x, \alpha)$ in the second argument on a given interval. Assumption 5.$(ii).b$-d guarantees that outside this interval, the value taken by $m(x, \alpha)$ is known. In particular, it imposes that variations of α are large enough to drive the willingness to overstay on the unit interval. Assumption 5.(iv) is needed to handle conditioning on probability zero events, such as $\{X_{i\bar{q}} = \bar{x}\}$. Assumption 5.$(iii)$ is a normalization given (i) and $(ii).a$. It implies that the following conditional distribution is identified from the data:

$$F_{\alpha_i|X_{iQ},X_{i\bar{q}}}(a|x,\bar{x}) = F_{P_{i\bar{q}}|X_{iQ},X_{i\bar{q}}}(a|x,\bar{x}) \tag{4.13}$$

Hence, by Assumption 5.$(i)-(iv)$, for all x such that $(X_{iQ}, X_{i\bar{q}}) = (x, \bar{x})$ and for all a such that $\underline{\alpha}(x) < a < \overline{\alpha}(x)$:

$$\begin{aligned} Q_{P_{iQ}|X_{iQ},X_{i\bar{q}}}\left(F_{P_{i\bar{q}}|X_{iQ},X_{i\bar{q}}}(a|x,\bar{x})|x,\bar{x}\right) &= Q_{m(X_{iQ},\alpha_i)|X_{iQ},X_{i\bar{q}}}\left(F_{P_{i\bar{q}}|X_{iQ},X_{i\bar{q}}}(a|x,\bar{x})|x,\bar{x}\right) \\ &= Q_{m(X_{iQ},\alpha_i)|X_{iQ},X_{i\bar{q}}}\left(F_{\alpha_i|X_{iQ},X_{i\bar{q}}}(a|x,\bar{x})|x,\bar{x}\right) \\ &= m\left(x, Q_{\alpha_i|X_{iQ},X_{i\bar{q}}}\left(F_{\alpha_i|X_{iQ},X_{i\bar{q}}}(a|x,\bar{x})|x,\bar{x}\right)\right) \\ &= m(x,a). \end{aligned} \tag{4.14}$$

Equation (4.14) makes it clear that two features of the data are key for nonparametric identification. First, continuity of the dependent variable P_{iQ} ensures the continuity of the quantile and cumulative distribution functions used in Equation (4.14). Second, repeated observation of the dependent variable in different hypothetical scenarios allows defining quantile and cumulative distribution functions for P_{iQ} and $P_{i\bar{q}}$.

Note that $m(x, a)$ is identified for all x, such that $(x, \bar{x}) \in Supp(X_{iQ}, X_{i\bar{q}})$. Hence, identification of $m(x, a)$ depends on the richness of the support $B(Q)$, for individual such that $B(\bar{q}) = \bar{y}$. The next assumption allows extrapolating on the whole support.

Assumption 6 (Analytic extrapolation)**.** *For each $q \in [0, 1]$, $Supp\left((B_i(q), q), (\bar{y}, \bar{q})\right)$ contains an interval and, for all $y \in Supp(B_i(q))$, the function $y \mapsto m(y, q, a)$ is real analytic on the real line.*

Under assumption 6, extrapolation is based on the property that real analytic functions that coincide on an open interval coincide everywhere. A similar condition is introduced in Arellano and Bonhomme (2017). The semiparametric model used in the empirical application satisfies Assumption 6.

Consider the identification of the conditional distribution of α. Note that for each x:

$$\underline{\alpha}(x) = \inf\{a : m(x,a) > 0\} \tag{4.15}$$

$$\overline{\alpha}(x) = \sup\{a : m(x,a) < 1\} \tag{4.16}$$

Furthermore, for all $P_{iQ} \in (0,1)$:

$$Q_{P_{iQ}|X_{iQ}}(q|x) = Q_{m(X_{iQ},\alpha)|X_{iQ}}(q|x) = m\left(x, Q_{\alpha_i|X_{iQ}}\right)$$

By the monotonicity condition 5.(ii):

$$\begin{cases} Q_{\alpha_i|X_{iQ}}(q|x) & = & m^{-1}\left(x, Q_{P_{iQ}|X_{iQ},X_{i\bar{q}}}(q|x)\right) \text{ if } P_{iQ} \in (0,1) \\ Q_{\alpha_i|X_{iQ}}(q|x) & < & \underline{\alpha}(x) \text{ if } P_{iQ} = 0 \\ Q_{\alpha_i|X_{iQ}}(q|x) & > & \overline{\alpha}(x) \text{ if } P_{iQ} = 1 \end{cases} \tag{4.17}$$

This completes the proof of identification.

Remark 1. *Assumption (ii).a. can be relaxed to allow for weakly increasing functions $a \mapsto m(x,a)$. In this case, α_i is only bounded, in the following way:*

$$\begin{aligned} LB(q,x) := \sup\left\{\alpha : m(X_{iQ},\alpha) \leq Q_{P_{iQ}|X_{iQ},X_{i\bar{q}}}(q|x)\right\} \\ < Q_{\alpha_i|X_{iQ}}(q|x) \leq \\ \inf\left\{\alpha : m(X_{iQ},\alpha) \geq Q_{P_{iQ}|X_{iQ},X_{i\bar{q}}}(q|x)\right\} := UB(q,x). \end{aligned} \tag{4.18}$$

If in addition, $m(x,a)$ is increasing in x:

$$\sup_{\tilde{x} \leq x} LB(q,\tilde{x}) < Q_{\alpha_i|X_{iQ}}(q|x) \leq \inf_{\tilde{x} \geq x} UB(q,\tilde{x}). \tag{4.19}$$

5 Estimation

The nonparametric estimation of the objects of interest comes with strong data requirements and the curse of dimensionality. In particular, estimation of the functions $m(x,a)$ and F_α requires estimation of the conditional CDFs of P_{iQ} and $P_{i\bar{q}}$ on the joint support of $(X_{iQ}, X_{i\bar{q}})$ (equations (4.14) and (4.17)). The present approach to circumventing these inherent difficulties in nonparametric estimation builds on Chernozhukov et al. (2020) to propose a semiparametric specification of the conditional CDF of P_{iQ} and $P_{i\bar{q}}$. The specification also alleviates the support requirements of Assumption 3 and Assumption 6 needed for nonparametric identification of the distribution parameters.

Assumption 7 (Parametric distribution)**.** *There exists some variable Z_i and a vector function $\pi(.)$ defined on the unit line such that: $F_{P_{iQ}|X_{iQ},X_{i\bar{q}},Z_i}(p|x) = \Gamma\left(R'_{iQ}\pi(p)\right)$ where $R_{iQ} = r(X_{iQ}, X_{i\bar{q}}, Z_i)$ and Γ is a known strictly increasing continuous CDF such as the standard normal or the logistic CDF.*

Assumption 7 is a flexible representation that allows non-separability in observed and unobserved characteristics. Z includes observable characteristics such as gender, age, city of residence and current legal status. It is important to note that, unlike

traditional approaches to models with subjective expectations, parametric assumptions are not introduced for the purpose of identification, but to alleviate strong data requirements for the purpose of estimation. The proposed estimation procedure consists in three steps:
Step 1. Estimate the conditional cumulative distribution function of P_{iQ}, for $Q \in (0,1)$ under Assumption 7.
Step 2. Obtain an estimate $\hat{\alpha}_i$.
Step 3. Estimate each object of interest based on the quantities estimated in Step 1 and Step 2.

Step 1

The estimation problem is similar to the Distribution Regression (DR) problem in Chernozhukov et al. (2020):

$$\hat{F}_{P_{iq}}(p|X_{iq}, X_{i\bar{q}}, Z_i) = \Gamma\left(R'_{iq}\hat{\pi}(p)\right),$$
$$\text{where } R_{iq} = r(X_{iQ}, X_{i\bar{q}}, Z_i),\ p \in [0,1),\ q \in \mathcal{Q}_i \text{ and}$$
$$\hat{\pi}(p) \in \arg\min_{\pi \in \mathbb{R}^{dim(R)}} \textstyle\sum_{q \in \mathcal{Q}_i} \sum_i 1\{R_{iq} \leq p\} \log(\Gamma(R'_i\pi)) + 1\{R_{iq} > p\} \log(1 - \Gamma(R'_{iq}\pi)) \tag{5.1}$$

As opposed to traditional estimation strategies, this step does not require adjustment for extreme observations.

Step 2

Instead of a direct estimation of the conditional CDF , it is simpler to estimate α.[8] Recall that:

$$P_{iQ} = m(X_{iQ}, \alpha_i), P_{iQ} \in (0,1)$$

where $m(x,a) = Q_{P_{iQ}|X_{iQ},X_{i\bar{q}}}\left(F_{P_{i\bar{q}}|X_{iQ},X_{i\bar{q}}}(a|x,\bar{x})|x,\bar{x}\right)$. By Assumption 5, $m(x,\alpha)$ is invertible in its second argument and it follows that:

$$\alpha_i = Q_{P_{i\bar{q}}|X_{iQ},X_{i\bar{q}}}\left(F_{P_{iQ}|X_{iQ},X_{i\bar{q}}}(P_{iQ}|x,\bar{x})|x,\bar{x}\right)$$

For a fine grid of the unit interval $[0,1]$ with S_p equidistant points, the following estimator is proposed:

$$\hat{\alpha}_i = \frac{1}{|\mathcal{Q}_i|} \sum_{Q \in \mathcal{Q}_i} \delta_p \sum_{s=1}^{S_p} \left(1 - 1\left\{\hat{F}_{P_{i\bar{q}}|X_{iQ},X_{i\bar{q}},Z_i}(p_s|X_{iQ},\bar{x},Z_i) \geq \hat{F}_{P_{iQ}|X_{iQ},X_{i\bar{q}},Z_i}(P_{iQ}|X_{iQ},\bar{x},Z_i)\right\}\right)$$

[8] As opposed to Evdokimov (2010) that cannot obtain an estimate of α_i because of the idiosyncratic disturbance U_{it}. Appendix D deals with the case where there exists an additional idiosyncratic disturbance that is interpreted as measurement error.

where $\delta_p = 1/S_p$ and $|\mathcal{Q}_i|$ is the number of elements (cardinal) of the set $\mathcal{Q}_i$ (typically four).[9]

The estimation of α takes advantage of the pseudo-panel structure that gives observations of P_{iQ} for at most four different values of Q. Although the estimator is not consistent for each individual, it can be used for consistent estimation of the object of interest. Note that about 130 respondents (16% of the sample) report $P_{iQ} = 1$ for all Q. For these individuals, $\hat{\alpha}$ would always equal one. Indeed, $\hat{F}_{P_{iQ}|X_{iQ},X_{i\bar{q}},Z_i}(1|X_{iQ},\bar{x},Z_i) = 1$. This would result in the distribution of costs and surplus being unbounded. One solution is to construct bounds on the CDFs for plausible values of a lower bound of costs and surpluses. The other solution adopted in the rest of the paper is to interpolate $\hat{F}_{P_{iq}}(p|X_{iq},X_{i\bar{q}},Z_i) = \Gamma\left(R'_{iq}\hat{\pi}(p)\right)$ for $p = 1$, using the estimated values for $p < 1$. Estimation results without the interpolation are available upon request.

Step 3

For any given value $y \in \mathbb{R}, \bar{q} \in [0,1], a \in [0,1]$, an estimator for $m(y,\bar{q},a))$ is given by the following expression:

$$\hat{m}(y,\bar{q},a) = \frac{1}{|\mathcal{Q}_i|}\sum_{Q\in\mathcal{Q}_i}\frac{1}{n}\sum_{i=1}^{n}\delta_p\sum_{s=1}^{S_p}\left(1 - 1\left\{\hat{F}_{P_{iQ}|X_{iQ},X_{i\bar{q}}}(p_s|y,\bar{q},\bar{x}) \geq \hat{F}_{P_{i\bar{q}}|X_{iQ},X_{i\bar{q}}}(a|y,\bar{q},\bar{x})\right\}\right). \tag{5.2}$$

Note that a direct estimation of $m(y,\bar{q},a)$, although feasible is not necessary for all objects of interest. Take for example, the CDF of the distribution of costs.

$$\begin{aligned}
\Pr\left(F_C(c;q,\alpha) \leq p\right) &= \int_0^1 1\left\{m(c,q,a) \leq p\right\} F_\alpha(da) \\
&= \int_0^1 1\left\{Q_{P_{iQ}|X_{iQ},X_{i\bar{q}}}\left(F_{P_{i\bar{q}}|X_{iQ},X_{i\bar{q}}}\left(a|(c,q),\bar{x}\right)|(c,q),\bar{x}\right) \leq p\right\} F_\alpha(da) \\
&= \int_0^1 1\left\{F_{P_{i\bar{q}}|X_{iQ},X_{i\bar{q}}}\left(a|(c,q),\bar{x}\right) \leq F_{P_{iQ}|X_{iQ},X_{i\bar{q}}}\left(p|(c,q),\bar{x}\right)\right\} F_\alpha(da)
\end{aligned}$$

Given the estimated quantities $\hat{F}_{P_{i\bar{q}}}$, $\hat{F}_{P_{iQ}}$, and $\hat{\alpha}_i$, the following estimator can be used:

$$\frac{1}{|\mathcal{Q}_i|}\sum_{Q\in\mathcal{Q}_i}\frac{1}{n}\sum_{i=1}^{n}1\left\{\hat{F}_{P_{i\bar{q}}|X_{iQ},X_{i1}}\left(\hat{\alpha}_i|(c,q),\bar{x}\right) \leq \hat{F}_{P_{iQ}|X_{iQ},X_{i1}}\left(p|(c,q),\bar{x}\right)\right\} \tag{5.3}$$

The derivation of estimators for the remaining objects of interest is presented in Appendix B.3. Inference is conducted using a weighted bootstrap procedure. The weighted bootstrap versions of each estimator can be obtained by rerunning the estimation procedure with sampling weights drawn from a distribution that satisfies Assumption 3 in Section 4 of

[9] For a bounded variable X, a consistent estimator of the τ-quantile of X is given by: $\delta_p\sum_{s=1}^{S_p}\left(1 - 1\left\{F_X \geq \tau\right\}\right)$. See Chernozhukov et al. (2020) for a similar estimator of the Quantile Structural Function.

Chernozhukov et al. (2020). The procedure for uniform inference is a simple adaptation of Algorithm 2 in Section 5.1 therein.

6 Empirical Analysis

This section presents the estimation results for the distributions of *ex ante* benefits, costs, surpluses and uncertainty in the population of Afghan asylum seekers. The section begins with a description of the sample, then provides some descriptive evidence about the intention to overstay in the population. Finally, it presents the estimation results.

6.1 Sample Characteristics

The data used for the empirical analysis stem from the "Survey on Migrants' expectations in Germany" that was conducted during the second half of 2019. The raw sample consists of 1,024 Afghan citizens, age 18 or more, who arrived in Germany for the first time in 2014 or after, and live in one of three urban areas: Berlin, Hamburg, or Munich, the German cities with the highest numbers of Afghan citizens. Computer assisted personal interviews were conducted by native speaker enumerators in Dari and Pashto, the two main languages spoken in Afghanistan. Recruitment was a mix of traditional sampling based on available register data, and peer-recruitment. Details of the survey implementation can be found in Méango et al. (2020).

Although not representative by design, the sample mimics key characteristics of the population. Table G.2 shows some demographic characteristics by city of residence. The sample is dominated by male. The population is young (median age 28), less educated than comparable Germans with nearly two-thirds of respondents with lower secondary education or below. The second half of Table G.2 presents additional characteristics related to the current stay in Germany. The average length of stay in Germany is 3.6 years. Nearly six of 10 respondents have received some form of protection, with this statistic varying considerably by city. The current level of employment is low and gender unequal (26% for men, 4.5% for women), as is participation in education in Germany (23% for men, 18% for women), which is dominated by vocational education. Most (80%) have been or are currently enrolled in a German language class. These sample characteristics are similar with those of the IAB-BAMF-SOEP survey (Brücker et al., 2018), a representative survey of asylum seekers in Germany.

6.2 Descriptive Evidence on Subjective Beliefs and Intention to Overstay

Table 6.1 presents averages and standard deviations of elicited subjective beliefs and intention to overstay for people in each city and the whole sample. For the subsequent analysis, the sample is restricted to respondents with non-missing answers in all key variables of interest, which is 858 observations. The expected income distributions are also trimmed at the top 95-percentile to minimize the effect of outliers. Respondents report on average a 68% chance of obtaining the RtS. Beliefs in Munich are significantly lower than in the two other cities with a 14 to 17 pp difference. In Berlin, the average belief is slightly above the 2019 official proportion of asylum seekers with some protection (72% compared to 68%), while it is lower in Hamburg (70% compared to 80%). For Munich, the only publicly available comparison point is at State level, Bavaria. The average belief of 58% is 12 pp lower compared to the proportion in Bavaria in 2019 (68%). However, the average belief is slightly higher than the proportion of Afghans who received positive decisions in the first instance (51%). Standard deviations are large, which suggests a significant variation between individual beliefs. Beliefs about the chance to obtain the RtS conditional on overstaying three years are slightly lower than, but well aligned with, the elicited chance to obtain the RtS.

Beliefs about the proportion of Afghans forcibly removed and sent back to Afghanistan and the chance of being deported when not obtaining the RtS are unreasonably high. On average, respondents believe that one out of five Afghans has been sent back to Afghanistan in the past years and there is a 37% chance of being deported conditional on not obtaining the RtS. As discussed in Section 2, deportation to Afghanistan is a rare event. In 2019, only 1.8 out of 1000 Afghan asylum seekers and 1.6 out of 100 Afghan asylum seekers with a rejected asylum application were deported.

With a RtS, respondents expect, on average, a monthly income of 1,676 EUR. This amount is lowest in Munich (1,610 EUR), which also displays the lowest variance, and highest in Hamburg (1,710 EUR). Without a RtS, respondents expect, on average, 1,261 EUR. As before, the average is lowest in Munich (1,122 EUR), and highest in Hamburg (1,424 EUR). These numbers imply an expected monthly return on legalization of between 300 EUR and 500 EUR on average, depending on the city.[10] The average expected income abroad is, on average, lower than the income expected in Germany (1,121 EUR), which implies, on average, a pecuniary return of 140 EUR from staying without a RtS. Note that the average expected income abroad is particularly low in the Munich sub-sample (815 EUR) compared to the other cities.

[10]According to Brücker et al. (2020), the average monthly gross income of refugees who entered Germany between 2013 and 2016 was 1,282 EUR in 2018, and 1,863 EUR for those in a full-time occupation. This represents between 54% and 89% of the average gross income of comparable German workforce, depending on the category considered.

	Berlin	Hamburg	Munich	Total sample
Obtain RtS (Q1)	74.29	70.24	58.36	69.22
	(25.90)	(24.04)	(26.30)	(26.34)
Obt. RtS. after 3 yrs w/o RtS (Q2)	67.12	67.74	52.31	63.55
	(27.64)	(25.94)	(26.30)	(27.64)
Be deported if no RtS (Q3)	30.31	37.43	48.01	36.55
	(31.85)	(26.77)	(24.14)	(29.69)
Income with RtS	1693.92	1710.70	1610.62	1676.36
	(745.1)	(565.5)	(511.4)	(648.3)
Income w/o RtS	1247.96	1424.31	1121.64	1260.99
	(726.5)	(558.3)	(557.3)	(650.5)
Income abroad	1217.68	1237.62	814.06	1121.45
	(868.4)	(611.4)	(799.5)	(812.3)
Stay w/o RtS (Q4)	71.84	69.90	49.88	65.82
	(30.88)	(26.28)	(27.91)	(30.46)
Stay w/o RtS after 3 yrs (Q5)	71.23	65.51	52.30	65.03
	(32.12)	(27.01)	(29.49)	(31.19)
stay if q=1% (Q6)	69.76	55.12	36.65	57.74
	(34.69)	(32.43)	(31.93)	(36.08)
stay if q=50% (Q7)	87.04	74.14	68.34	79.08
	(21.26)	(22.30)	(24.90)	(23.90)
stay if q=99% (Q8)	96.65	89.88	92.43	93.88
	(10.57)	(16.30)	(15.89)	(13.92)

Note: Mean values calculated on non-missing observations. Standard deviation in parentheses.

Table 6.1: Sample characteristics by city

Intention to overstay as elicited through Q4 is high with an average of 64%. One of four respondents state a 100% chance to overstay, of whom close to 80% reside in Berlin. The intention to overstay in Munich is markedly lower than in the two other cities, with a 21 pp difference, on average.

Intention to overstay as elicited with questions Q6 to Q8 is also high. When $q = 0.99$, it displays an average of 93.23. The CDFs, in this case, are highly skewed as 68% of the respondents answer 100, and close to 90% give answers of 75 or above. This pattern is very consistent across all cities (see Figure G.2 in Appendix G). The average intention to overstay drops by about 15 pp when $q = 0.50$. The magnitude of this change depends strongly on the city: Berlin, -10 pp; Hamburg, -16 pp; Munich, -25 pp. The distribution is more spread: 42% of the sample answer 100, and close to 90% give answers of 45 or above. When $q = 0.01$, the drop in the average intention to overstay from the case where $q = 0.99$ is, on average, 35 pp. Again, the magnitude of this change depends strongly on the city: Berlin, -26 pp; Hamburg, -44 pp; Munich, -54 pp. Thus, Munich residents appear the least willing to stay when there is almost no chance of becoming regularized. Overall, the chance of obtaining the RtS in the future appears to have a significant effect on the intention to overstay.

6.3 Estimated Distributions

6.3.1 Pecuniary benefits

The expected pecuniary benefits $B(q)$ can be directly calculated from elicited observations given Equation (A.4) in Appendix A. The yearly discount factor is set to 0.95 and the number of sub-periods T to 36 months. Pecuniary benefits, costs, and surpluses are all expressed in monthly amount equivalents over a six-year period. The lifetime-equivalent of these estimates depend on the time horizon considered. However, Appendix A shows that the relative importance of each component does not.

Figure 6.1 presents the cumulative distribution at different values of q, 0.10, 0.50, 0.90, and at the elicited value Q_i. When $q = q_i$ (solid black line), the bulk of expected pecuniary benefits is between 0 and 500 EUR, with the first quartile being 0 and the third quartile 412 EUR. The CDFs have a negative tail on the left; about 17% of the observations show a negative expected pecuniary benefit from overstaying, while 95% of the population expect benefits below 1,000 EUR monthly. Variations in Q moves the CDF to the right. The increase is, average, 48 EUR from $q = 0.10$ to $q = 0.50$, and a further 48 EUR from $q = 0.50$ to $q = 0.90$.

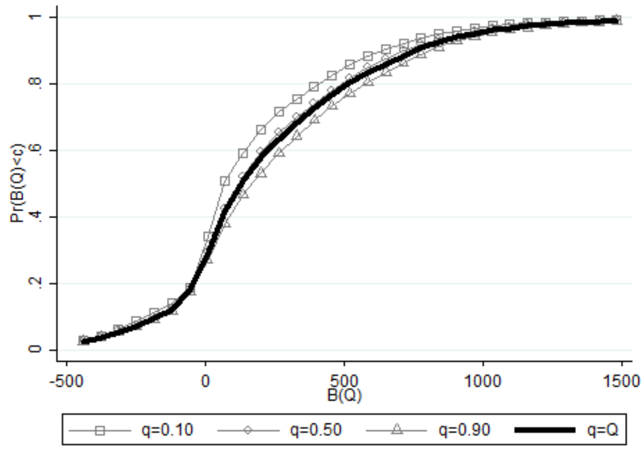


Figure 6.1: Distribution of $B(q)$: pecuniary benefits

6.3.2 Costs

The estimation strategy detailed in Section 5 is applied to the data with the following choices for Equation (5.1):

$$\bar{y} = 0, \quad \bar{q} = 0.01, \tag{6.1}$$

$$S_p = 100, \text{ i.e. } p = 0, 0.01, 0.02, \cdots, 0.99. \tag{6.2}$$

$$\Gamma(x) = \exp(x)/(1 + \exp(x)) \tag{6.3}$$

$$r(X_{iQ}, X_{i\bar{q}}, Z_i) = \theta_1(B_i(Q_i)) + \theta_2(B_i(\bar{q})) + Q \times \text{City } \beta'_{\text{City}} + Z\gamma' \tag{6.4}$$

where $B_i(q)$ is the calculated pecuniary benefit at q, City is an indicator variable for the respondent's city of residence and Z includes individual characteristics (gender, age, education, city of residence, length of stay in Germany and current legal status). Experimenting with several specifications, linear functions $\theta_1(.)$ and $\theta_2(.)$ are found to work reasonably well.

The exposition of the results related to distributional objects of interest requires some extra care. The objects of interest are CDFs on the population. However, each individual holds a different distribution of beliefs about costs and surpluses that depend on α. Thus, we are interested in CDFs of CDFs.

Figure 6.2 presents the CDF $\Pr(\Pr(C(\bar{q}, \alpha, \nu) \leq c) \geq p)$, for different values of p (0.10, 0.50, 0.90) and q (0.10, 0.50, 0.90, q_i). The shaded areas represent the 90% confidence interval for the value $q = q_i$ based on 200 bootstrap replications. The CDFs represent the proportion of the population that assigns a probability of at least p to the event "the cost of overstating is lower than a given value c." Consider for example the extreme case $p = 0.90$ (Figure 6.2(a)). It represents the proportion of the population that assigns a probability of at least 0.90, hence a large probability, to the event that the cost of overstaying is lower than a given value c, with c varying on the x-axis between -1500 and

2500 EUR.

Given current subjective beliefs ($q = q_i$, solid black line), a quarter of the population expects negative costs of overstaying with a large probability, and 60% expects costs below 1,000 EUR with a large probability. These proportions appear very sensitive to subjective beliefs about the chance of becoming regularized q. If the population held common belief $q = 0.10$, less than 5% of the population would assign a large probability to the event that the cost is negative, and only a quarter would expect costs below 1,000 EUR with a large probability. By contrast, if the population held common belief $q = 0.90$, more than 40% of the population would assign a large probability to the event of negative costs. As much as 85% would expect costs below 1,000 EUR with a large probability.

At the other extreme ($p = 0.10$), Figure 6.2.(c) corresponds to the case where individuals assign at least some non-negligible probability. It shows that virtually all the population assigns some non-negligible weights to the event of negative costs.

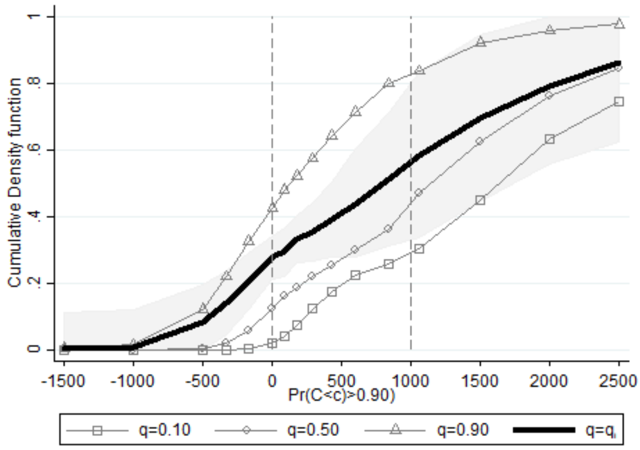
Finally, Figure 6.2(b) shows the case where $p = 0.50$. Conveniently, under Assumption 4, it also corresponds to the CDF of the average *ex ante* cost $\bar{C}(q, \alpha)$. The following exposition uses this interpretation. Given current subjective beliefs ($q = q_i$, solid black line), close to 70% of the population expects negative costs on average, and virtually all the population expects costs below 1,000 EUR on average. The CDF moves to the left with increases in the chance of becoming regularized q. If the population held the common belief that $q = 0.10$, only 20% of the population would expect negative costs on average. If q were 0.50, this proportion would increase to 60%, and to 95% when $q = 0.90$.

In summary, the expected costs of overstaying are relatively low in the population, but depend crucially on the expected chance of becoming regularized.

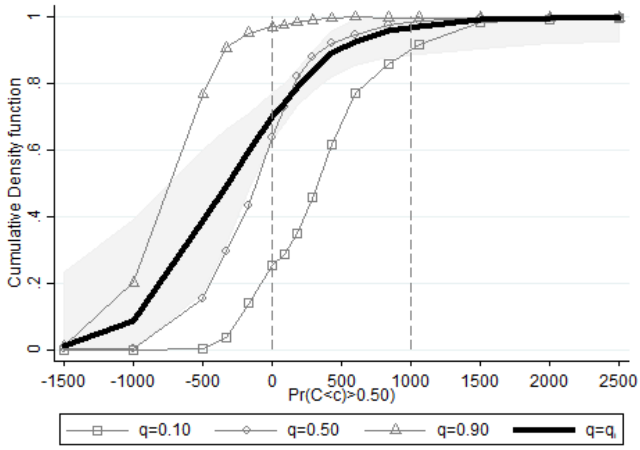
6.3.3 Surpluses

Figure 6.3 presents similar results for the distribution of *ex ante* surpluses and is a mirror image of Figure 6.2. Figure 6.3(a) shows that irrespective of the chance of becoming regularized, virtually no one in the population expects a negative surplus with great probability. Only when $q = 0.10$ does one-fifth of the population expect surpluses lower than 2,000 EUR with great probability. Nevertheless, the majority of the population assigns non-negligible weights to the possibility of negative surpluses, in particular when $q = 0.10$ (see Figure 6.3(c)).

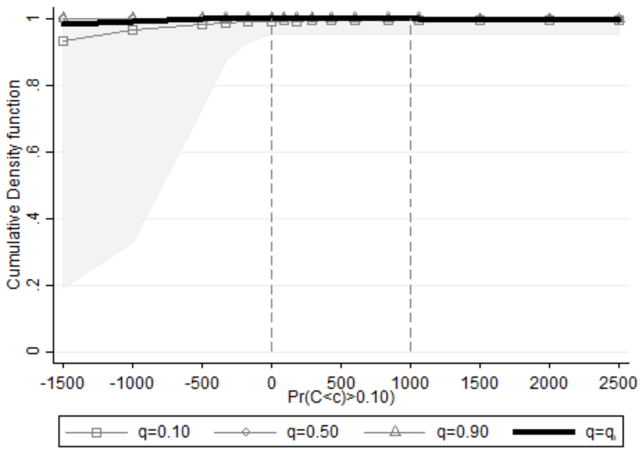
Figure 6.3(b) can be interpreted as the CDF of the average surplus $\bar{S}(b(q), q, a)$ under Assumption 4. It shows that at current beliefs, about 20% of the population expects negative average surpluses, and about 60% of the population expects surpluses between 0 and 1,000 EUR. The lower-middle quartile expects surpluses between 0 and 600 EUR, and the upper-middle quartile between 600 and 900 EUR. The CDF moves to the right with q, as higher surpluses are expected with a better chance of becoming regularized.



(a) $p = 0.90$

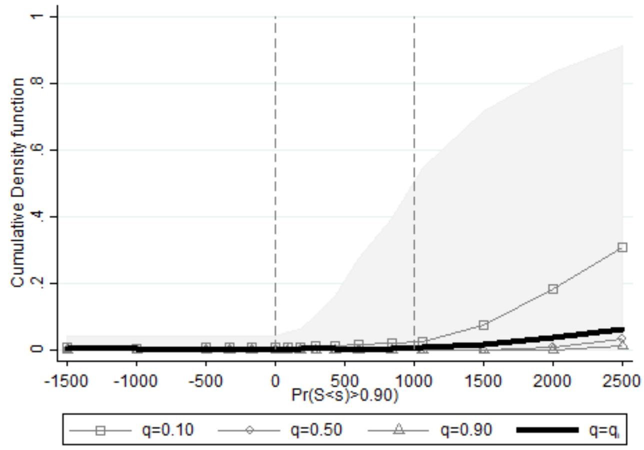


(b) $p = 0.50$

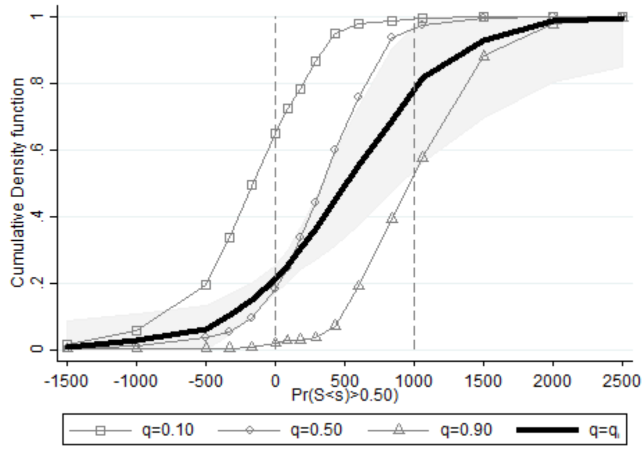


(c) $p = 0.10$

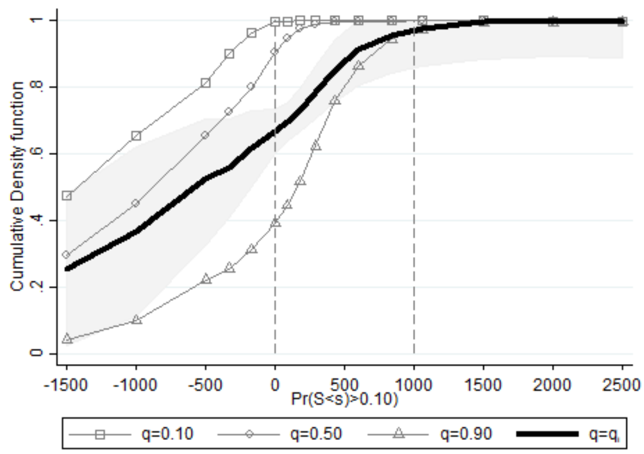
Figure 6.2: Distribution of *ex ante* costs $\Pr\left(\Pr\left(C(q,\alpha,\nu) \leq c\right) \geq p\right)$



(a) $p = 0.90$



(b) $p = 0.50$



(c) $p = 0.10$

Figure 6.3: Distribution of *ex ante* surpluses $\Pr\left(\Pr\left(S(b(q), q, \alpha, \nu) \leq s | \alpha, b(q)\right) \geq p\right)$

Average surpluses are strictly negative for more than 60% of the population when $q = 0.10$ and strictly positive for almost all the population when $q = 0.90$.

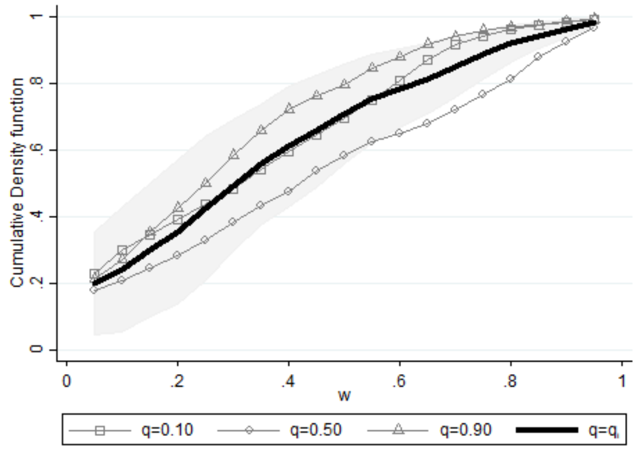


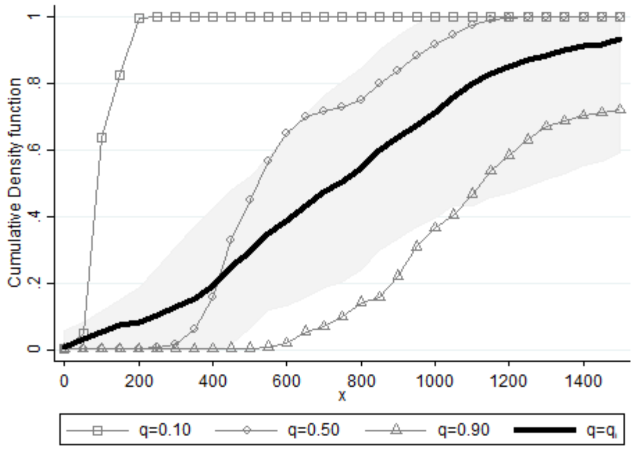
Figure 6.4: Distribution of W_B, share of pecuniary benefits in average surpluses

It is instructive to calculate the share of pecuniary benefits in the *ex ante* average surplus as defined by Equation (B.16). An estimator of its CDF is given in Appendix B.3. Figure 6.4 presents the estimated CDF for the usual values of q. The x-axis represents the share of average surpluses explained by pecuniary benefits. Consider the distribution of current beliefs (solid black line). For about one-quarter of the population, pecuniary benefits represent less than 10% of the expected surplus. This proportion is about 30% for the median individual and reaches 50% around the third quartile. An increase chance of becoming regularized has a non-monotonic effect on the share of surplus explained by pecuniary benefits. The CDF first moves to the left as q increases from 0.10 to 0.50, then to the right as int further increases to 0.90. The variation are however not substantial. In summary, for most of the population, pecuniary benefits explain only a modest part of expected surpluses.

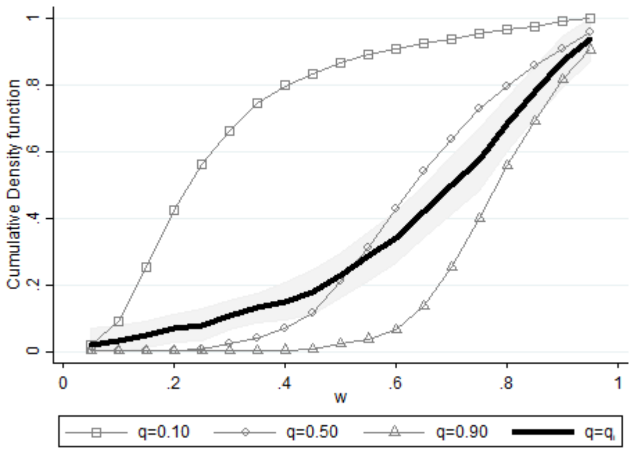
6.3.4 Option-Value

The option-value (OV) related to the option to be regularized, as defined in Equation (B.11), provides another measure of the importance of the chance of becoming regularized. Figure 6.5 presents the CDF of the calculated OV and the CDF of the share of the OV in the expected surplus as defined by Equation (B.14) for the usual values of q.

The OV plays an important role, being close to 500 EUR at the first quartile, which is as much as 50% of average surpluses. The OV is about 700 EUR for the median individual, which is as much as 70% of average surpluses. The OV explains more than 80% of average surpluses for the top quartile. Not surprisingly, the importance of the OV increases noticeably with q. Note that part of the OV is generated by pecuniary benefits through the difference $B(q) - B(0)$. However, this quantity represents a small part of



(a) $\Pr\left(OV(q,\alpha) \leq x\right)$



(b) $\Pr\left(W_{OV}(q,\alpha) \leq w\right)$

Figure 6.5: Distribution of option value

the OV (64 EUR, on average, for $q = q_i$). These findings highlight the importance of amenities related to the RtS.

6.3.5 Distribution of Resolvable Uncertainty

Figure 6.6 displays the unconditional and conditional CDF of the resolvable uncertainty parameter ν at selected values α, under the assumption of separability.

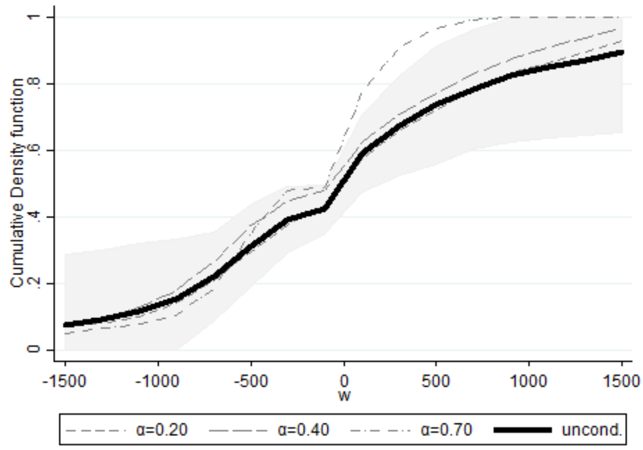


Figure 6.6: Distribution of uncertainty F_ν and $F_{\nu|\alpha}$

The unconditional distribution is quite spread with an interquartile range close to 1,300 EUR. This amounts to 1.5 times the interquartile range of the average *ex ante* surplus distribution. Hence, uncertainty is relatively important in the population. Note that the distribution varies with α, with the variance decreasing with increasing α. This fact warrants the use of the present estimation strategy, which relaxes the assumption of common beliefs in the population.

6.3.6 Elasticities

The elasticity of the probability of emigration with respect to income and the chance of becoming regularized are of particular interest.[11] Their magnitude sheds some light on the effect of policies that would induce a marginal change, either in income or access to the RtS. A one-percent decrease from the average pecuniary benefits (227 EUR) results, on average, in a decrease by 0.06% (p-value$<$ 0.10) of the intention to overstay. In contrast, a one-percent decrease from the average chance of becoming regularized decreases, on average, by 0.39% (p-value$<$ 0.01) the intention to overstay. These findings confirm the modest role played by pecuniary benefits and the importance of the chance of becoming regularized.

[11]The reported average elasticity with respect to the variable X is calculated as: $\mathbb{E}(\dfrac{\partial P_Q}{\partial X})/\dfrac{\bar{P}_Q}{\bar{X}}$

7 A Social Planner's Problem

This section takes the perspective of a social planner (SP) who wishes to minimize host-country's social costs associated with overstaying migrants. Given the importance of the chance of becoming regularized, q, in the asylum seeker's decision, the discussion focuses on a SP's problem, which consists in choosing an optimal value q.[12]

Assume that the mean cost associated with the presence of an asylum seeker who overstays without the RtS is C_I. This cost may consist of social transfers, as explained in Section 2. It may also reflect a social aversion to irregular stay.[13] Similarly, denote by C_R the mean cost associated with the presence of an asylum seeker who overstays and obtains the RtS. This costs may also consist of social transfers, although in the case of regularized migrants, these transfers are expected to be lower because of a full access to the labor market (see, for example, Orrenius and Zavodny, 2012; Devillanova et al., 2018). The mean cost also reflects the labor market costs for some native workers (see, for example, Monras et al., 2018), the (lack of) public support for regularizing overstayers, or a social preference for some type of migrants (see, for example, Bansak et al., 2016). In the following, it is assumed that $C_I > C_R \geq 0$.

Setting q at some value q_0 has several consequences. First, it changes the number of asylum seekers with a rejected application who overstay: on average, a proportion $\mathbb{E}m\left(B(q_0), q_0, \alpha\right)$ decides to stay. Second, it divides the population of overstayers into two groups: those who become regularized, a proportion q_0, with associated mean cost C_R, and those who do not, a proportion $(1 - q_0)$, with associated mean cost C_I. For simplicity, it is assumed that the regularization decision is random. Third, regularization creates a so-called "pull-effect". Each new regularization attracts a proportion π of new migrants. These new migrants are attracted by the prospect of regularization. Therefore, they would not qualify as humanitarian migrants. For example, when $\pi = 0.25$, four regularized overstayers attract one additional migrant. Finally, asylum seekers' *ex ante* investments in country-specific human capital, for example, language acquisition, reacts to the prospect of regularization. Therefore, a high probability of becoming regularized reduces the cost associated parameter C_I and C_R by a factor φ.[14]

[12]Using the estimate of the surplus distribution, one could also consider a policy that pays financial incentives in exchange for a voluntary return. However, given the low elasticity of the intention to overstay with respect to pecuniary benefits, this policy will only have a moderate effect.

[13]For example, Wang (2012) comments: "Although the weight of evidence suggests that immigration is not linked to crime, the public consistently views immigrants, especially undocumented immigrants, as criminal and thus a threat to social order."

[14]Mukhopadhyay (2019) finds a link between the probability of deportation and the education decision of illegal migrants. Khourshed and Méango (2020) show that Syrian refugees in Germany who expect higher chance of obtaining a permanent resident status from German language acquisition are more likely to invest in it.

The SP's problem can be written as follows:

$$\min_q (1 - \varphi q) \cdot (1 + \sum_{t=1,2,\cdots} (\pi q)^t) \cdot (C_R q + C_I(1 - q)) \cdot \mathbb{E}m\,(B(q), q, \alpha) \qquad (7.1)$$

where the term $\sum_{t=1,2,\cdots}(\pi q)^t)$ reflects the fact that newly regularized migrants attract further migration. The social cost depends on the parameters π, φ, and C_R/C_I.

Figure G.4 in Appendix G represents the cost function for different parameter values $(\pi, C_R/C_I)$, $\varphi = 0$, and q ranging from 0.1 to 0.9. The cost function is concave in q, first increasing, then decreasing, implying that minimum-cost policies are at either one of the two extremes. Hence, the solution to the SP's problem is either a "restrictive regularization" policy (small q) that minimizes the number of overstayers or a "large-scale regularization" (large q) that takes advantage of a small cost C_R. The optimal policy depends on the parameters π, φ, and C_R/C_I.

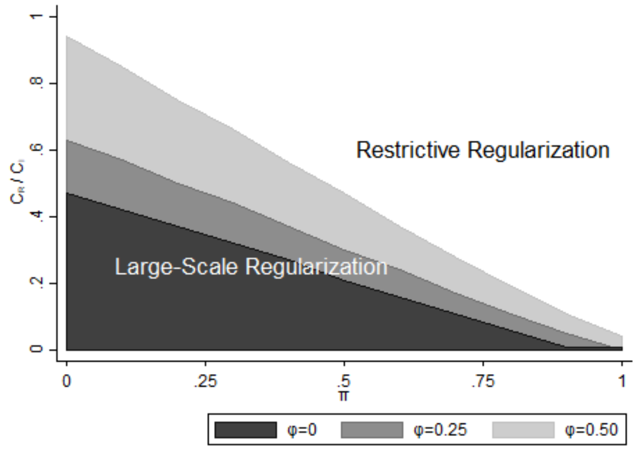


Figure 7.1: Optimal choice

Figure 7.1 illustrates the optimal policy choice as a function of different parameter combinations. The x-axis represents the "pull-effect" parameter π. The y-axis represents the mean-cost ratio C_R/C_I. The shaded areas represent regions where a large-scale regularization is preferable to a restrictive policy. The (overlayed) areas represent different values of φ.

When either or both C_R/C_I and π are small, a large-scale regularization is preferable to restrictive regularization policies. More specifically, without "pull-effect", $\pi = 0$, a sufficient condition for a large-scale regularization to be optimal is that C_R is lower than 0.45 C_I. Conversely, if regularized overstayers do not generate additional social costs, C_R/C_I close to 0, a sufficient condition is that the pull-effect factor is smaller than 0.8. A large-scale regularization can remain optimal even in the case where both parameters π and C_R/C_I take moderate values.

The size of the shaded region increases with φ, implying that a large-scale regularization becomes optimal for additional pairs $(\pi, C_R/C_I)$. When φ is large, human capital investments are very elastic to the chance of becoming regularized. A large-scale regularization becomes optimal by reducing both costs associated with regularized and non-regularized overstayers.

In summary, a social planner who wishes to minimize the cost associated with overstaying asylum seekers has the choice between one of two policies: restrictive regularization or large-scale regularization. Policies in-between are sub-optimal. A restrictive regularization policy saves costs by deterring asylum seekers from overstaying. However, the intention to overstay is high on average, even when the chance of becoming regularized is low. A sufficient condition for a large-scale regularization to be cost-efficient is that social costs associated with regularized migrants and pull-effect of regularization are small to moderate. Moreover, when costs associated with regularized migrants are small, a large-scale regularization remains preferable to a restrictive policy even if the pull-effect is large. Finally, when asylum seekers' investments in country-specific human capital increase with better chance of regularization, a large-scale regularization becomes optimal for more combinations of social costs and pull-effect.[15]

8 Discussion

The objective of this paper was to calculate asylum seekers' *ex ante* returns on overstaying. *Ex ante* surpluses of overstaying are positive for the majority of the population, but very heterogeneous: about 20% of the population expects, on average, negative surpluses of overstaying. These findings agree with the high average intention to stay expressed by asylum seekers and with the large variance in these intentions. Moreover, the results highlight the modest contribution of pecuniary benefits in the *ex ante* surplus. In contrast, the option-value created by the chance of becoming regularized appears as a key determinant of the *ex ante* returns. Afghan asylum seekers are ready to spend a long time with a precarious status to eventually obtain the right to stay and the amenities associated with it.

This paper has several contributions to the literature. The first set of contributions is methodological. It proposes a decision model for asylum seekers' choice of overstaying. It provides a novel identification strategy of several objects of interest for models using choice probabilities. Simultaneously, it builds on the existing literature to offer a feasible semiparametric estimation strategy. The second set of contributions is empirical. The paper measures mostly positive *ex ante* returns on overstaying, although the risk of deportation is perceived as very high. It shows that beliefs about the chance of becoming regularized are the main driver of these returns. Given the importance of those beliefs,

[15] Casarico et al. (2018) provide a related discussion on the desirability of amnesty programs.

further research should be conducted to understand how asylum seekers form their expectations about the different risks they face while overstaying, and to what extent those beliefs are malleable. Finally, the findings highlight the importance of amenities related to the legal right to stay.

The structural model could easily be extended to include risk aversion. However, this would complicate the identification of distributional objects of interest because pecuniary benefits and cost function would no longer be separable. Nevertheless, the lack of separability would not affect the estimation of the elasticity parameters, which only depend on the reduced-form representation of the intention to stay. The estimated elasticities can then be used to recover the option-value parameter, as well as expected surpluses up to normalization.

The analysis in this paper puts into perspective the consequences of a political strategy of deterrence (in German, *Abschreckung*), which aims at decreasing future opportunities for a regular stay to avoid creating so-called pull-effects. The intention to overstay of those already present is high. Therefore, the cost-saving effect of deterrence is limited. Given that deportation is rarely enforced, it is likely that a large part of the asylum seekers will remain in Germany, irrespective of their status. Furthermore, there is a lack of evidence for a sizable effect of regularization programs on subsequent immigration flows (see, for example, Orrenius and Zavodny, 2003; Wong and Kosnac, 2017; Cascio and Lewis, 2020). Moreover, there is growing evidence that regularized migrants increase fiscal revenues, thanks to a better access to the labor market (see, for example, Monras et al., 2018, and references therein). Finally, a low prospect of regularization might deter asylum seekers *ex ante* from human capital investments that are key for their integration. In this context, a large-scale regularization should be considered as a potential cost-efficient policy.

References

Amemiya, T. (1985). *Advanced econometrics*. Harvard university press.

Arellano, M. and Bonhomme, S. (2017). Quantile selection models with an application to understanding changes in wage inequality. *Econometrica*, 85(1):1–28.

Attanasio, O. P. and Kaufmann, K. M. (2014). Education choices and returns to schooling: Mothers' and youths' subjective expectations and their role by gender. *Journal of Development Economics*, 109:203–216.

Bah, T. L. and Batista, C. (2018). Understanding willingness to migrate illegally: Evidence from a lab in the field experiment. Working Paper No. 1803. Universidade Nova de Lisboa, Faculdade de Economia, NOVAFRICA.

Bansak, K., Hainmueller, J., and Hangartner, D. (2016). How economic, humanitarian, and religious concerns shape European attitudes toward asylum seekers. *Science*, 354(6309):217–222.

Bayer, P., Khan, S., and Timmins, C. (2011). Nonparametric identification and estimation in a Roy model with common nonpecuniary returns. *Journal of Business & Economic Statistics*, 29(2):201–215.

Blass, A. A., Lach, S., and Manski, C. F. (2010). Using elicited choice probabilities to estimate random utility models: Preferences for electricity reliability. *International Economic Review*, 51(2):421–440.

Borjas, G. J. (1987). Self-selection and the earnings of immigrants. *The American Economic Review*, 77:531–553.

Brücker, H., Kosyakova, Y., and Schuß, E. (2020). Fünf jahre seit der Fluchtmigration 2015: Integration in Arbeitsmarkt und Bildungssystem macht weitere Fortschritte. Technical report, IAB-Kurzbericht 04/2020.

Brücker, H., Rother, N., and Schupp, J. (2018). *IAB-BAMF-SOEP-Befragung von Geflüchteten 2016: Studiendesign, Feldergebnisse sowie Analysen zu schulischer wie beruflicher Qualifikation, Sprachkenntnissen sowie kognitiven Potenzialen*, volume 30. DEU.

Casarico, A., Facchini, G., and Frattini, T. (2018). What drives the legalization of immigrants? evidence from IRCA. *Regional Science and Urban Economics*, 70:258–273.

Cascio, E. U. and Lewis, E. G. (2020). Opening the door: Migration and self-selection in a restrictive legal immigration regime. Technical report, National Bureau of Economic Research.

Chernozhukov, V., Fernández-Val, I., Newey, W., Stouli, S., and Vella, F. (2020). Semiparametric estimation of structural functions in nonseparable triangular models. *Quantitative Economics*, 11(2):503–533.

Chernozhukov, V. and Hansen, C. (2005). An IV model of quantile treatment effects. *Econometrica*, 73(1):245–261.

Delavande, A. (2008). Pill, patch, or shot? Subjective expectations and birth control choice. *International Economic Review*, 49(3):999–1042.

Delavande, A. and Kohler, H.-P. (2016). Hiv/aids-related expectations and risky sexual behaviour in Malawi. *The Review of Economic Studies*, 83(1):118–164.

Devillanova, C., Fasani, F., and Frattini, T. (2018). Employment of undocumented immigrants and the prospect of legal status: evidence from an amnesty program. *ILR Review*, 71(4):853–881.

d'Haultfoeuille, X. and Maurel, A. (2013). Inference on an extended Roy model, with an application to schooling decisions in France. *Journal of Econometrics*, 174(2):95–106.

Eisenhauer, P., Heckman, J. J., and Vytlacil, E. (2015). The generalized Roy model and the cost-benefit analysis of social programs. *Journal of Political Economy*, 123(2):413–443.

Evdokimov, K. (2010). Identification and estimation of a nonparametric panel data model with unobserved heterogeneity. *Working Paper, Princeton University.*

Florens, J.-P., Heckman, J. J., Meghir, C., and Vytlacil, E. (2008). Identification of treatment effects using control functions in models with continuous, endogenous treatment and heterogeneous effects. *Econometrica*, 76(5):1191–1206.

Grogger, J. and Hanson, G. H. (2011). Income maximization and the selection and sorting of international migrants. *Journal of Development Economics*, 95(1):42–57.

Heckman, J. J. (2001). Micro data, heterogeneity, and the evaluation of public policy: Nobel lecture. *Journal of Political Economy*, 109(4):673–748.

Heckman, J. J. and Vytlacil, E. (2005). Structural equations, treatment effects, and econometric policy evaluation 1. *Econometrica*, 73(3):669–738.

Henry, M., Méango, R., and Mourifié, I. (2020). Revealing gender-specific costs of stem in an extended Roy model of major choice. Unpublished Manuscript.

Hoxhaj, R. (2015). Wage expectations of illegal immigrants: The role of networks and previous migration experience. *International Economics*, 142:136–151.

Imbens, G. W. and Newey, W. K. (2009). Identification and estimation of triangular simultaneous equations models without additivity. *Econometrica*, 77(5):1481–1512.

Jensen, R. (2010). The (perceived) returns to education and the demand for schooling. *The Quarterly Journal of Economics*, 125(2):515–548.

Khourshed, M. and Méango, R. (2020). Language proficiency and economic incentives: The case of Syrian asylum seekers in Germany. MEA Discussion Paper 01-2020.

Mbaye, L. M. (2014). "barcelona or die": Understanding illegal migration from Senegal. *IZA Journal of Migration*, 3(1):21.

McKenzie, D., Gibson, J., and Stillman, S. (2013). A land of milk and honey with streets paved with gold: Do emigrants have over-optimistic expectations about incomes abroad? *Journal of Development Economics*, 102:116–127.

Méango, R., Khourshed, M., and López-Falcón, D. (2020). From asylum seekers to illegal migrants: The intention to overstay of Afghan asylum seekers. Unpublished Manuscript.

Miller, G., De Paula, Á., and Valente, C. (2020). Subjective expectations and demand for contraception. Technical report, National Bureau of Economic Research. NBER Working Paper 27271.

Monras, J., Vázquez-Grenno, J., and Ferran, E. (2018). Understanding the effects of legalizing undocumented immigrants. CEPR Discussion Paper No. DP12726.

Mukhopadhyay, S. (2019). Legal status and immigrants' educational investment decisions. *Review of Economics of the Household*, 17(1):1–29.

Orrenius, P. M. and Zavodny, M. (2003). Do amnesty programs reduce undocumented immigration? Evidence from IRCA. *Demography*, 40(3):437–450.

Orrenius, P. M. and Zavodny, M. (2012). The economic consequences of amnesty for unauthorized immigrants. *Cato J.*, 32:85.

Sjaastad, L. A. (1962). The costs and returns of human migration. *Journal of Political Economy*, 70(5, Part 2):80–93.

Stinebrickner, R. and Stinebrickner, T. R. (2014). A major in science? Initial beliefs and final outcomes for college major and dropout. *Review of Economic Studies*, 81(1):426–472.

Van der Klaauw, W. (2012). On the use of expectations data in estimating structural dynamic choice models. *Journal of Labor Economics*, 30(3):521–554.

Wang, X. (2012). Undocumented immigrants as perceived criminal threat: A test of the minority threat perspective. *Criminology*, 50(3):743–776.

Wiswall, M. and Zafar, B. (2015). Determinants of college major choice: Identification using an information experiment. *The Review of Economic Studies*, 82(2):791–824.

Wong, T. K. and Kosnac, H. (2017). Does the legalization of undocumented immigrants in the us encourage unauthorized immigration from Mexico? An empirical analysis of the moral hazard of legalization. *International Migration*, 55(2):159–173.

Appendix

A Extended Model

This section presents an extension of the model to include a risk of deportation within each period and discusses the possibility of including a longer time horizon. In contrast, the simplified model of Section 3 has two periods and a risk of deportation only at the end of the first period.

Let λ_i be the perceived probability to be deported in each month, with $(1 - p_i^D) = (1 - \lambda_i)^T$. Denote by β the monthly discount factor. Expected gains in period 1 from staying in Germany without a RtS:

$$\sum_{t=1}^{T} \beta^t \left[(1-\lambda_i)^t \cdot \left(\alpha_i^{N,G} - \alpha_i^E + \gamma_i \left(Y_i^{N,G} - Y_i^E\right)\right) - c_i^D \cdot (1-\lambda_i)^{t-1} \cdot \lambda_i\right]$$

Expected gains in period 2 from staying in Germany without a RtS and obtaining subsequently a RtS in period 2:

$$\beta^T \cdot (1 - p_i^D) \cdot Q_i \left[\sum_{t=T+1}^{2T} \beta^{t-T} \left(\alpha_i^{R,G} - \alpha_i^E + \gamma_i \left(Y_i^{R,G} - Y_i^E\right)\right)\right]$$

Expected gains in period 2 from staying in Germany without a RtS and <u>not</u> obtaining subsequently a RtS in period 2:

$$\beta^T \cdot (1 - p_i^D) \cdot (1 - Q_i) \cdot P_i^{t=1}$$
$$\left[\sum_{t=T+1}^{2T} \beta^{t-T} \cdot \left((1-\lambda_i)^t \left(\alpha_i^{N,G} - \alpha_i^E + \gamma_i \left(Y_i^{N,G} - Y_i^E\right)\right) - c_i^D \cdot (1-\lambda_i)^{t-1} \cdot \lambda_i\right)\right]$$

Thus, with previous notations:

$$\begin{aligned}
\zeta_{0i} &= \sum_{t=1}^{T}\beta^t\left[c_i^D\cdot(1-\lambda_i)^{t-1}\cdot\lambda_i-(1-\lambda_i)^t\cdot\left(\alpha_i^{N,G}-\alpha_i^E\right)\right] && \text{(A.1)}\\
&+\beta^T\cdot(1-p_i^D)\cdot P_i^{t=1}\\
&\left[\sum_{t=T+1}^{2T}\beta^{t-T}\cdot\left(c_i^D\cdot(1-\lambda_i)^{t-1}\cdot\lambda_i-(1-\lambda_i)^t\left(\alpha_i^{N,G}-\alpha_i^E\right)\right)\right]\\
\zeta_{1i} &= \beta^T\cdot(1-p_i^D)\left[-\sum_{t=T+1}^{2T}\beta^{t-T}\left(\alpha_i^{R,G}-\alpha_i^E\right)+\right. && \text{(A.2)}\\
&\left.+P_i^{t=1}\sum_{t=T+1}^{2T}\beta^{t-T}\cdot\left(c_i^D\cdot(1-\lambda_i)^{t-1}\cdot\lambda_i-(1-\lambda_i)^t\left(\alpha_i^{N,G}-\alpha_i^E\right)\right)\right]\\
B_i(Q_i) &= \left(Y_i^{R,G}-Y_i^E\right)\beta^T\cdot(1-p_i^D)\cdot Q_i\sum_{t=T+1}^{2T}\beta^{t-T} && \text{(A.3)}\\
&+\left(Y_i^{N,G}-Y_i^E\right)\left[\sum_{t=1}^{T}\beta^t\cdot(1-\lambda_i)^t+\beta^T\cdot(1-p_i^D)\cdot(1-Q_i)\cdot P_i^{t=1}\sum_{t=T+1}^{2T}\beta^{t-T}\cdot(1-\lambda_i)^t\right]
\end{aligned}$$

The model considered so far two periods of three years. A possible strategy to extends the time horizon is to assume that individual repeats the game for S three-year periods, with consistent beliefs about the chance of deportation, p_i^D, the chance to obtain the RtS, Q_i, and the chance to exit in the next period, $P_i^{t=1}$. Costs and benefits parameters are easy to adapt in this case. For example:

$$\begin{aligned}
B_i(Q_i) &= \left(Y_i^{R,G}-Y_i^E\right)\sum_{s=2}^{S}\left[\left(\beta^T\cdot(1-p_i^D)\right)^{(s-1)}(1-Q_i)^{(s-2)}\cdot Q_i\sum_{t=(s-1)T+1}^{ST}\beta^{t-(s-1)T}\right] && \text{(A.4)}\\
&+\left(Y_i^{N,G}-Y_i^E\right)\sum_{s=1}^{S}\left[\left(\beta^T\cdot(1-p_i^D)\cdot(1-Q_i)\cdot P_i^{t=1}\right)^{(s-1)}\sum_{t=(s-1)T+1}^{sT}(\beta\cdot(1-\lambda_i))^{t-(s-1)T}\right]
\end{aligned}$$

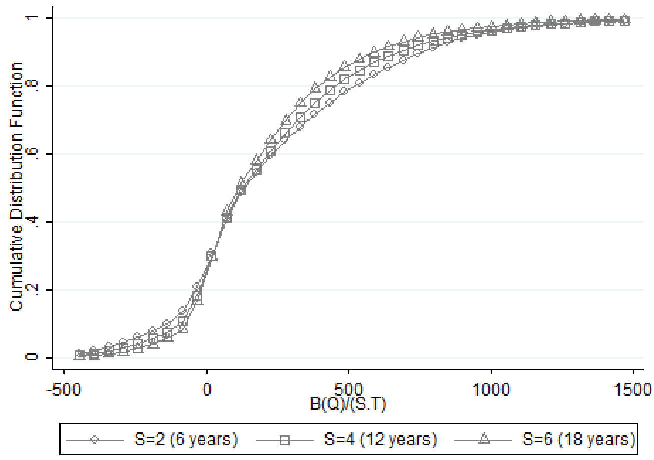


Figure A.1: $B_i(Q_i)/(S\times T)$ for $S=2,4,6$ and $T=36$

Figure A.1 shows that calculation of a monthly equivalent $B_i(Q_i)/(S \times T)$ produces a distribution of pecuniary benefits, which depends little on S. Therefore, the results on the relative importance of pecuniary benefits and costs are not affected by the time horizon chosen for the analysis.

B Proofs

B.1 Proof of Equation (4.5)

Under assumptions 1 and 2, Equation (4.1) can be rewritten:

$$
\begin{aligned}
P_{iQ} &= \int 1\left\{S_{iQ} \geq 0\right\} F_{\nu_i|\alpha_i}\left(d\nu|\alpha_i\right) \\
&= \int 1\left\{\bar{S}(B_i(Q_i), Q_i, \alpha_i) + V_{iQ} \geq 0\right\} F_{\nu_i|\alpha_i}\left(d\nu|\alpha_i\right) \\
&= \int 1\left\{V_{iQ} \geq -\bar{S}(B_i(Q_i), Q_i, \alpha_i)\right\} F_{\nu_i|\alpha_i}\left(d\nu|B_i(Q_i), Q_i, \alpha_i\right) \\
&= \int_{\mathcal{V}(Q)} 1\left\{v \geq -\bar{S}(B_i(Q_i), Q_i, \alpha_i)\right\} F_{V_{iQ}|\alpha_i}\left(dv|\alpha_i\right) \\
&= \int_{\mathcal{V}(Q)} 1\left\{v \geq -\bar{S}(B_i(Q_i), Q_i, \alpha_i)\right\} F_{V_{iQ}|\alpha_i}\left(dv|B_i(Q_i), Q_i, \alpha_i\right) \\
&= 1 - F_{V_{iQ}|Q_i,\alpha_i}\left(-\bar{S}(B_i(Q_i), Q_i, \alpha_i)|Q_i, \alpha_i\right)
\end{aligned} \tag{B.1}
$$

where the fourth line uses assumption 2. Indeed, under this assumption, there is a one-to-one mapping between ν and V_Q at given values of Q and α.

B.2 Objects of interest expressed as functionals $\Lambda\left(m(., q, a\right)$

B.2.1 Distributions of cost and surplus

One can derive the individual-specific distribution of *ex ante* surpluses for a given individual with private information α, given q and $b(q)$:

$$
\begin{aligned}
F_S\left(s; b(q), q, a\right) &= \Pr\left(S(b(q), q, \alpha, \nu) \leq s|\alpha = a, b(q)\right) \\
&= \int I\left\{b(q) - C(q, a, \nu) \leq s\right\} F_{\nu|\alpha, B(q)}(d\nu|a, b(q)) \\
&= \int I\left\{b(q) - C(q, a, \nu) \leq s\right\} F_{\nu|\alpha}(d\nu|a) \\
&= \int I\left\{(b(q) - s) - C(q, a, \nu) \leq 0\right\} F_{\nu|B(Q), Q, \alpha}(d\nu|b(q) - s, q, a) \\
&= \int I\left\{B(Q) - C(Q, \alpha, \nu) \leq 0\right\} F_{\nu|B(Q), Q, \alpha}(d\nu|b(q) - s, q, a) \\
&= 1 - m(b(q) - s, q, a)
\end{aligned}
$$

The second line uses the independence condition ensuring that $F_{\nu|\alpha}(\nu|a) = F_{\nu|B(Q),Q,\alpha}(\nu|y - s, q, a)$. And the last line follows from the model (4.2).

Given knowledge of the distributions, one can derive the first moments as in Chernozhukov et al. (2020).

$$\begin{aligned}\bar{C}(q,a) &= \int_{\mathcal{C}^+} [1 - F_C(c;q,a)]\,dc - \int_{\mathcal{C}^-} [F_C(c;q,a)]\,dc \\ &= \int_{\mathcal{C}^+} [1 - m(c,q,a)]\,dc - \int_{\mathcal{C}^-} [m(c,q,a)]\,dc \end{aligned} \tag{B.2}$$

$$\bar{S}(b(q),q,a) = b(q) - \bar{C}(q,a) \tag{B.3}$$

B.2.2 Derivatives of cost and surplus

Identification of the above quantity requires the support condition from Assumption 3. By contrast, the quantities below do not require such condition. Instead, they require either separability of the resolvable uncertainty parameter or the strict monotonicity of Assumption 2. The derivative of the emigration probability is given by:

$$\frac{\partial P_Q}{\partial q} = \nabla_q m(B(Q),Q,\alpha). \tag{B.4}$$

Consider the derivative of surplus. Under Assumption 1 and Assumption 2, taking the derivative of Equation (4.5) with respect to q yields:

$$\frac{\partial P_Q}{\partial q} = \frac{\partial \bar{S}(B(Q),Q,\alpha)}{\partial q} \cdot F'_{V|Q,\alpha}\left(\bar{S}(B(Q),Q,\alpha)|Q,\alpha\right).$$

It suffices to note that:

$$\begin{aligned}\nabla_1 m(B(Q),Q,\alpha) &= \frac{\partial\left(F_{V|Q,\alpha}\left(\bar{S}(B(Q),Q,\alpha)|Q,\alpha\right)\right)}{\partial b(q)} \\ &= \frac{\partial \bar{S}(B(Q),Q,\alpha)}{\partial b(q)} \cdot F'_{V|Q,\alpha}\left(\bar{S}(B(Q),Q,\alpha)|Q,\alpha\right)\end{aligned} \tag{B.5}$$

and $\dfrac{\partial \bar{S}(B(Q),Q,\alpha)}{\partial b(q)} = 1$, to conclude that:

$$\frac{\partial \bar{S}(B(Q),Q,\alpha)}{\partial q} = \frac{\nabla_q m(B(Q),Q,\alpha)}{\nabla_1 m(B(Q),Q,\alpha)} \tag{B.6}$$

The derivative of expected *ex ante* cost follows easily as:

$$\frac{\partial \bar{C}(Q,\alpha)}{\partial q} = \frac{\nabla_q m(B(Q),Q,\alpha)}{\nabla_1 m(B(Q),Q,\alpha)} - \frac{\partial B(Q)}{\partial q} \tag{B.7}$$

Note that upon knowledge of the average surplus, Equation (B.5) provides an expression of the density of the resolvable uncertainty.

$$F'_{V|Q,\alpha}\left(-\bar{S}(B(Q),Q,\alpha)|Q,\alpha\right) = \nabla_1 m(B(Q),Q,\alpha)$$

Without Assumption 3, $\bar{C}$ and $\bar{S}$ are only identified up to an additive function of q and α.

B.2.3 Option-value

The option value is given by integrating the derivative of the average surplus:

$$\bar{S}(b(q),q,\alpha) - \bar{S}(b(0),0,\alpha) = \int_0^q \frac{\nabla_q m(b(q),q,\alpha)}{\nabla_1 m(b(q),q,\alpha)}dq \tag{B.8}$$

B.3 Derivation of estimators

Distribution of *ex ante* surpluses

$$\begin{aligned}
&\Pr\left(F_S(s;B(q),q,\alpha)\leq p\right)\\
&= \textstyle\int_0^1\int_{\mathcal{B}(q)} 1\left\{1-m(y-s,q,a)\leq p\right\} F_{B(q|\alpha)}(dy|a)F_\alpha(da)\\
&= \textstyle\int_0^1\int_{\mathcal{B}(q)} 1\left\{m(y-s,q,a)\geq 1-p\right\} F_{B(q|\alpha)}(dy|a)F_\alpha(da)\\
&= \textstyle\int_0^1\int_{\mathcal{B}(q)} 1\left\{Q_{P_{iQ}|X_{iQ},X_{i\bar{q}}}(F_{P_{i\bar{q}}|X_{iQ},X_{i1}}\left(a|(y-s,q),\bar{x}\right)|(y-s,q),\bar{x})\geq 1-p\right\} F_{B(q)}(dy|a)F_\alpha(da)\\
&= \textstyle\int_0^1\int_{\mathcal{B}(q)} 1\left\{F_{P_{i\bar{q}}|X_{iQ},X_{i1}}\left(a|(y-s,q),\bar{x}\right)\geq F_{P_{iQ}|X_{iQ},X_{i1}}\left(1-p|(y-s,q),\bar{x}\right)\right\} F_{B(q)}(dy|a)F_\alpha(da)
\end{aligned} \tag{B.9}$$

Estimator:

$$\frac{1}{|\mathcal{Q}_i|}\sum_{Q\in\mathcal{Q}_i}\frac{1}{n}\sum_{i=1}^{n} 1\left\{\hat{F}_{P_{i\bar{q}}|X_{iQ},X_{i1}}\left(\hat{\alpha}_i|(y-s,q),\bar{x}\right)\geq \hat{F}_{P_{iQ}|X_{iQ},X_{i1}}\left(1-p|(y-s,q),\bar{x}\right)\right\} \tag{B.10}$$

Distribution of the average *ex ante* cost Recall that:

$$\begin{aligned}
\bar{C}(q,a) &= \int_{\mathcal{C}^+}\left[1-F_C(c;q,a)\right]dc - \int_{\mathcal{C}^-}\left[F_C(c;q,a)\right]dc\\
&= \int_{\mathcal{C}^+}\left[1-m(c,q,a)\right]dc - \int_{\mathcal{C}^-}\left[m(c,q,a)\right]dc
\end{aligned}$$

Hence:

$$\Pr\left(\bar{C}(q,\alpha)\leq\bar{c}\right) = \int_0^1 1\left\{\int_{\mathcal{C}^+}\left[1-m(c,q,a)\right]dc - \int_{\mathcal{C}^-}\left[m(c,q,a)\right]dc\leq\bar{c}\right\}F_\alpha(da)$$

Estimator:

$$\begin{aligned}
\hat{P}\left(\bar{C}(q,\alpha)\le\bar{c}\right) &= \frac{1}{|\mathcal{Q}_i|}\sum_{Q\in\mathcal{Q}_i}\frac{1}{n}\sum_{i=1}^{n}1\left\{\hat{\bar{C}}(q,\hat{\alpha}_i)\le\bar{c}\right\}\\
\hat{\bar{C}}(q,a) &= \delta_c\sum_{s=1}^{S_c}\left(1\left\{y_s\ge 0\right\}-\hat{m}(y_s,q,a)\right)\\
\hat{m}(y,q,a) &= \delta_p\sum_{s=1}^{S_p}\left(1-1\left\{\hat{F}_{P_{iQ}|X_{iQ},X_{i\bar{q}}}(p_s|y,q,\bar{x})\ge\hat{F}_{P_{i\bar{q}}|X_{iQ},X_{i\bar{q}}}(a|y,q,\bar{x})\right\}\right)
\end{aligned}$$

Distribution of the average cost under Assumption 4

$$\begin{aligned}
\bar{C}(q,a) &= \sup\{y:m(y,q,a)\le p_0\}\\
&= \sup\{y:Q_{P_{iQ}|X_{iQ},X_{i\bar{q}}}\left(F_{P_{i\bar{q}}|X_{iQ},X_{i1}}\left(a|(y,q),\bar{x}\right)|(y,q),\bar{x}\right)\le p_0\}\\
&= \sup\{y:F_{P_{i\bar{q}}|X_{iQ},X_{i1}}\left(a|(y,q),\bar{x}\right)\le F_{P_{iQ}|X_{iQ},X_{i\bar{q}}}\left(p_0|(y,q),\bar{x}\right)\}
\end{aligned}$$

Distribution of the option-value Recall that:

$$OV(q,\alpha):=\bar{S}(b(q),q,\alpha)-\bar{S}(b(0),0,\alpha) \tag{B.11}$$

$$\begin{aligned}
\Pr\left(OV(q,\alpha)\le x\right) &= \Pr\left(B(q)-\bar{C}(q,\alpha)-(B(0)-\bar{C}(0,\alpha))\le x\right)\\
&= \Pr\left(\bar{C}(q,\alpha)\ge B(q)-(B(0)-\bar{C}(0,\alpha))-x\right)\\
&= \Pr\left(p_0\ge m\left(B(q)-(B(0)-\bar{C}(0,\alpha))-x,q,\alpha\right)\right)\\
&= \Pr\left(F_{P_{i\bar{q}}|X_{iQ},X_{i1}}\left(\alpha|(x^\star,q),\bar{x}\right)\le F_{P_{iQ}|X_{iQ},X_{i\bar{q}}}\left(p_0|(x^\star,q),\bar{x}\right)\right)
\end{aligned} \tag{B.12}$$

where $x^\star=B(q)-(B(0)-\bar{C}(0,\alpha))-x$. The third line uses Assumption 4.
Estimator:

$$\frac{1}{|\mathcal{Q}_i|}\sum_{Q\in\mathcal{Q}_i}\frac{1}{n}\sum_{i=1}^{n}1\left\{\hat{F}_{P_{i\bar{q}}|X_{iQ},X_{i1}}\left(\hat{\alpha}|(x^\star,q),\bar{x}\right)\le\hat{F}_{P_{iQ}|X_{iQ},X_{i\bar{q}}}\left(p_0|(x^\star,q),\bar{x}\right)\right\} \tag{B.13}$$

Share of the option-value in the average *ex ante* surplus Define a measure for the share of the option value in the average *ex ante* surplus as follows:

$$W_{OV}(q,\alpha)=\frac{|\bar{S}(B(q),q,\alpha)-\bar{S}(B(0),0,\alpha)|}{|\bar{S}(B(0),0,\alpha)|+|\bar{S}(B(q),q,\alpha)-\bar{S}(B(0),0,\alpha)|},\ w\in(0,1). \tag{B.14}$$

$$\begin{aligned}
&\Pr\left(W_{OV}(q,\alpha)\le w\right)\\
&=\Pr\left(|\bar{S}(B(q),q,\alpha)-\bar{S}(B(0),0,\alpha)|\le|\bar{S}(B(0),0,\alpha)|\cdot\left(\frac{1}{w}-1\right)^{-1}\right)
\end{aligned} \tag{B.15}$$

As $\bar{S}$ increases with q, $\bar{S}(B(q),q,\alpha)-\bar{S}(B(0),0,\alpha)>0$. To obtain an estimator of the quantity of interest, it suffices to replace x with $|\bar{S}(B(0),0,\alpha)|\cdot\left(\frac{1}{w}-1\right)^{-1}$ in the expression of $x^\star$ (Equation (B.12)).

Share of pecuniary benefits in the average *ex ante* surpluses Define a measure for the share of pecuniary benefits in the average *ex ante* surplus as follows

$$W_B(q,\alpha) = \frac{|B(q)|}{|B(q)| + |\bar{S}(B(q),q,\alpha) - B(q)|},\ w \in (0,1). \tag{B.16}$$

$$\begin{aligned}
&\Pr\left(W_B(q,\alpha) \le w\right) \\
&= \Pr\left(|\bar{S}(B(q),q,\alpha) - B(q)| \ge |B(q)| \cdot \left(\frac{1}{w} - 1\right)\right) \\
&= 1 - \Pr\left(\bar{S}(B(q),q,\alpha) - \bar{S}(B(0),0,\alpha) \le |B(q)| \cdot \left(\frac{1}{w} - 1\right) + B(q) - \bar{S}(B(0),0,\alpha)\right) \\
&\quad + \Pr\left(\bar{S}(B(q),q,\alpha) - \bar{S}(B(0),0,\alpha) \le -|B(q)| \cdot \left(\frac{1}{w} - 1\right) + B(q) - \bar{S}(B(0),0,\alpha)\right)
\end{aligned} \tag{B.17}$$

An estimator is obtained in a similar fashion as in the previous paragraph.

SUPPLEMENTARY MATERIAL

For online publication only

C Sources and additional of Official statistics

Sources of official statistics accessible online, all last accessed on October 15,2020:

1. Eurostat: First instance decisions on applications by citizenship, age and sex - annual aggregated data (rounded). URL: https://ec.europa.eu/eurostat/databrowser/view/MIGR_ASYDCFSTA__custom_55039/default/table?lang=en.

2. Statistisches Bundesamt, DESTATIS, URL: https://www-genesis.destatis.de/genesis/online

    - Code 12531-0008: Persons seeking protection: Germany, reference date, sex, category of protection status/protection status, country groups/citizenship.
    - Code 12531-0026: Persons seeking protection: Länder, reference date, sex, category of protection status/protection status, country groups/citizenship.

3. Deportations and departure statistics from the federal government:

    - 2014 - Deutscher Bundestag, Drucksache 18/4025. URL: http://dipbt.bundestag.de/extrakt/ba/WP18/649/64916.html
    - 2015 - Deutscher Bundestag, Drucksache 18/7588. URL: http://dipbt.bundestag.de/extrakt/ba/WP18/717/71788.html
    - 2016 - Deutscher Bundestag, Drucksache 18/11112. URL: http://dipbt.bundestag.de/extrakt/ba/WP18/794/79434.html
    - 2017 - Deutscher Bundestag, Drucksache 19/800. URL: http://dipbt.bundestag.de/extrakt/ba/WP19/2312/231225.html
    - 2018 - Deutscher Bundestag, Drucksache 19/8201. URL: http://dipbt.bundestag.de/extrakt/ba/WP19/2436/243665.html
    - 2019 - Deutscher Bundestag, Drucksache 19/18201. URL: http://dipbt.bundestag.de/extrakt/ba/WP19/2589/258926.html

4. Compiled statistics on deportations by origin country, state (*Länder*) responsible of the deportation and year of deportation:
   Bundeszentrale für Poltische Bildung, URL: https://www.bpb.de/gesellschaft/migration/flucht/zahlen-zu-asyl/265765/abschiebungen-in-deutschland.

5. Short explanation of the toleration status:
   Bundeszentrale für Poltische Bildung, URL: https://www.bpb.de/gesellschaft/migration/kurzdossiers/233846/definition-fuer-duldung-und-verbundene-rechte?p=all

D Model with measurement error

This sections considers a measurement error on the left hand side. Suppose that the following is observed:

$$
\begin{aligned}
P_{iq} &= m(B_i(q_i), q_i, \alpha_i) + U_{iq} \text{ for some } q_i \in \{0, 0.01, 0.02, \ldots, 1\}. \\
P_{iq_{0j}} &= m(B_i(q_{0j}), q_{0j}, \alpha_i) + U_{ij}, \text{ for } q_{01} = 0.01, q_{02} = 0.50 \text{ and } q_{03} = 0.99.
\end{aligned}
$$

Assume for now that $P_{iQ} \in (0,1)$, for $Q \in \mathcal{Q}_i$. The identification results builds on Theorem 2 in Evdokimov (2010). Assume that:

Proposition 1. *(i) There exists a vector of observable characteristics Z (not including Q) such that the conditional probability distribution function of U_{iQ} obeys:*

$$f_{U_{iQ}|X_{iQ},\alpha_i,X_{i(-Q)},U_{i(-Q)}}(u|x,\alpha,x_{-q},u_(-q)) = f_{U_{iQ}|Z_i}(u|z) \tag{D.1}$$

and

$$f_{U_{it}|X_{it},\alpha_i,X_{i(-t)},U_{i(-t)}}(u|x,\alpha,x_{-t},u_(-t)) = f_{U_{it}|Z_i}(u|z) \tag{D.2}$$

for all Q and t.

(ii) $E(U_{iQ}|Z_i = z) = E(U_{it}|Z_i = z) = 0$ for all Q and z.

(iii) $\phi_{U_{iQ}}(s|Z_i = z)$ and $\phi_{U_{it}}(s|Z_i = z)$ do not vanish for all z, Q and t, where $\phi_Y(s|X = x)$ is the conditional characteristic function of Y.

(iv) $m(x, \alpha)$ is strictly increasing in α.

(v) α_i is continuously distributed conditional on X_{iQ}.

(vi) $m(x, \alpha)$, $f_{U_{it}|Z_i}(u|z)$, $f_{\alpha_i|X_{iQ}}(a|x)$, and $f_{\alpha_i|X_{iQ},X_{i\bar{q}}}(a|x_1, x_2)$ are everywhere continuous with respect to x, x_1, x_2, for all u and a.

(vii) there exists $\bar{x} = (\bar{y}, \bar{q})$, such that for all α $m(x, \alpha) = \alpha$.

then $m(x, a)$ and the conditional distribution of α_i are identified (on part of their support).

Remark 2. *In (i), we could allow the distribution of U_i to change with Q. However, it has to be identifiable from a finite number of points Q. For example $U_{iQ} = Z_i\beta + \gamma Q + V_{iQ}$ and $V_{iQ} \perp X_{iQ}$.*

The result of Evdokimov (2010) implies that distribution of U_{iQ} is identified from the joint distribution of P_{iQ} and $P_{i\bar{q}}$. Note that:

$$\begin{aligned}\phi_{P_{iQ}}(s|X_{iQ}=x,X_{i\bar{q}}=\bar{x}) &= \phi_{m(X_{iQ},\alpha_i)}(s|X_{iQ}=x,X_{i\bar{q}}=\bar{x})\phi_{U_{iQ}}(s|X_{iQ}=x)\\ \phi_{P_{i\bar{q}}}(s|X_{iQ}=x,X_{i\bar{q}}=\bar{x}) &= \phi_{\alpha_i}(s|X_{iQ}=x,X_{i\bar{q}}=\bar{x})\phi_{U_{iQ}}(s|X_{i\bar{q}}=\bar{x})\end{aligned}$$

The second line uses the normalization (vii).

Hence, the following conditional characteristic function of $m(X_{iQ},\alpha_i)$ is also identified as:

$$\phi_{m(X_{iQ},\alpha_i)}(s|X_{iQ}=x,X_{i\bar{q}}=\bar{x}) = \frac{\phi_{P_{iQ}}(s|X_{iQ}=x,X_{i\bar{q}}=\bar{x})}{\phi_{U_{iQ}}(s|X_{iQ}=x)} \tag{D.3}$$

and

$$\phi_{\alpha_i}(s|X_{iQ}=x,X_{i\bar{q}}=\bar{x}) = \frac{\phi_{P_{iQ}}(s|X_{iQ}=x,X_{i\bar{q}}=\bar{x})}{\phi_{U_{i\bar{q}}}(s|X_{i\bar{q}}=\bar{x})} \tag{D.4}$$

Furthermore, by (iv) - (vi), for all x such that $(X_{iQ},X_{i\bar{q}})=(x,\bar{x})$:

$$\begin{aligned}Q_{m(X_{iQ},\alpha_i)|X_{iQ},X_{i\bar{q}}}\left(F_{\alpha_i|X_{iQ},X_{i\bar{q}}}(a|x,\bar{x})\right) &= m\left(x,Q_{\alpha_i|X_{iQ},X_{i\bar{q}}}\left(F_{\alpha_i|X_{iQ},X_{i\bar{q}}}(a|x,\bar{x})|x,\bar{x}\right)\right)\\ &= m(x,a).\end{aligned}$$

Finally, the following conditional characteristic function of $m(X_{iQ},\alpha_i)$ is also identified:

$$\phi_{m(X_{iQ},\alpha_i)}(s|X_{iQ}=x) = \frac{\phi_{P_{iQ}}(s|X_{iQ}=x)}{\phi_{U_{iQ}}(s|X_{iQ}=x)} \tag{D.5}$$

and by (iv) - (vi):

$$Q_{\alpha_i|X_{iQ}}(q|x) = m^{-1}\left(x,Q_{m(X_{iQ},\alpha_i)|X_{iQ},X_{i\bar{q}}}(q|x)\right) \tag{D.6}$$

Remark 3 (Censoring)**.** *A significant proportion of the sample answers with extreme probabilities, in particular with intention to stay equals to 1. The framework would therefore fail because the measurement error would not be separable. The model can be extended to the following:*

$$\begin{aligned}P^*_{iQ} &= min(m(B_i(Q_i),Q_i,\alpha_i)+U_{iQ},1),\ \textit{for}\ Q\in\{0,0.01,0.02,\ldots,1\}\\ P^*_{iq_{0j}} &= min(m(B_i(q_{0j}),q_{0j},\alpha_i)+U_{iq_{0j}},1),\ \textit{for}\ q_{01}=0.01,q_{02}=0.50\ \textit{and}\ q_{03}=0.99.\end{aligned}$$

Thus, the data is censored above. The identification result follows easily, if the joint distribution of the truncated random variable $(P_{iQ},P_{i\bar{q}})$ can be recovered nonparametrically from the censored data.

E Random effect

The proposed estimator assumes that α may be correlated with $B(Q)$ and Q. This assumption can be relaxed in studies where respondents are presented with alternative scenario, within which outcome values are determined exogenously (see e.g. Blass et al., 2010; Bah and Batista, 2018). In this case, α can be seen as a random effect. This section addresses estimation under this framework.

Assumption 8 (parametric distribution)**.** *There exists some variable Z_i and τ_i such that:*

(i) $\alpha_i = h(Z_i, \tau_i)$ *with* $\tau_i \perp X_{iQ}|Z_i$ *and* $\tau_i|Z_i \sim \mathcal{U}[0,1]$, *where h is a continuous function, strictly increasing in its second argument.*

(ii) $F_{P_{iQ}|X_{iQ},Z_i}(p|x) = \Gamma\left(R_{iQ}'\pi(p)\right)$ *where* $R_{iQ} = r(X_{iQ}, Z_i)$ *and Γ is a known strictly increasing continuous CDF such as the standard normal or the logistic CDF.*

The first part of the assumption implies that α_i depends on observable characteristics and of a random effect τ_i that is independent of X_{iQ} and Z_i. The second part specifies a parametric distribution for the conditional P_{iQ}. It is a flexible representation that allows non-separability in observed and unobserved characteristics.

Assumption 8 considerably simplifies the identification result. Indeed, there is some τ such that:

$$m(x,\alpha) = m\left(x, h(z,\tau)\right) = \widetilde{m}(x,z,\tau) \tag{E.1}$$

Because $\tau_i \sim \mathcal{U}[0,1]$:

$$\begin{aligned}
Q_{P_{iQ}|X_{iQ},Z_i}(\tau|x,z) &= Q_{m(X_{iQ},\alpha_i)|X_{iQ},Z_i}(q|x,z) \\
&= Q_{\widetilde{m}(X_{iQ},Z_i,\tau_i)|X_{iQ},Z_i}(\tau|x,z) \\
&= \widetilde{m}\left(x,z,Q_{\tau_i|X_{iQ},Z_i}(\tau|x,z)\right) \\
&= \widetilde{m}\left(x,z,\tau\right)
\end{aligned}$$

The distribution of any element of interest $\Lambda(m(.,q,a))$ is then given by:

$$\int 1\left\{\Lambda(m(.,q,a)) \leq l\right\} F_\alpha(da) = \int_0^1 \int_{\mathcal{Z}} \left\{\Lambda(\widetilde{m}(.,q,z,\tau)) \leq l\right\} F_Z(dz)d\tau \tag{E.2}$$

The estimation problem is similar to DR problem in Chernozhukov et al. (2020):

$$\hat{F}_{P_{iQ}}(p|X_{iQ},Z_i) = \Gamma\left(R_i'\hat{\pi}(p)\right), \text{ where } R_i = r(X_{iQ},Z_i) \tag{E.3}$$

$$\hat{\pi}(p) \in \arg\min_{\pi\in\mathbb{R}^{dim(R)}} \sum_i e_i 1\{R_i \leq p\}\log(\Gamma(R_i'\pi)) + 1\{R_i > p\}\log(1-\Gamma(R_i'\pi)) \tag{E.4}$$

The following paragraphs present some examples of corresponding estimators.

Distribution of *ex ante* costs Denote by $F_C(c;q,a) := \Pr\left(C(q,\alpha,\nu) \leq c | \alpha = a\right)$

$$\begin{aligned}
\Pr\left(F_C(c,q,\alpha) \leq p\right) &= \int 1\left\{m(c,q,a) \leq p\right\} F_\alpha(da) \\
&= \int_0^1 \int_{\mathcal{Z}} 1\left\{\widetilde{m}(c,q,z,\tau) \leq p\right\} F_Z(dz)d\tau \\
&= \int_0^1 \int_{\mathcal{Z}} 1\left\{Q_{P_{iQ}|X_{iQ},Z_i}(\tau|c,q,z) \leq p\right\} F_Z(dz)d\tau \\
&= \int_{\mathcal{Z}} F_{P_{iQ}|X_{iQ},Z_i}(p|c,q,z)F_Z(dz)
\end{aligned}$$

Estimator:

$$\frac{1}{|\mathcal{Q}_i|}\sum_{Q\in\mathcal{Q}_i}\frac{1}{n}\sum_{i=1}^n \hat{F}_{P_{iQ}|X_{iQ},Z_i}(p|c,q,Z_i) \tag{E.5}$$

Distribution of *ex ante* surpluses

$$\begin{aligned}
\Pr\left(F_S(s;B(q),q,\alpha) \leq p\right) &= \int\int_{\mathcal{B}(q)} 1\left\{1 - m(y-s,q,a) \leq p\right\} F_{B(q|\alpha)}(dy|a)F_\alpha(da) \\
&= \int_0^1 \int_{\mathcal{Z}}\int_{\mathcal{B}(q)} 1\left\{\widetilde{m}(y-s,q,z,\tau) \geq 1-p\right\} F_{B(q|Z)}(dy|z)F_Z(dz)d\tau \\
&= \int_0^1 \int_{\mathcal{Z}}\int_{\mathcal{B}(q)} 1\left\{Q_{P_{iQ}|X_{iQ},Z_i}(\tau|y-s,q,z) \geq 1-p\right\} F_{B(q|Z)}(dy|z)F_Z(dz)d\tau \\
&= 1 - \int_{\mathcal{Z}}\int_{\mathcal{B}(q)} F_{P_{iQ}|X_{iQ},Z_i}(1-p|y-s,q,z)F_{B(q|Z)}(dy|z)F_Z(dz)
\end{aligned}$$

Estimator:

$$1 - \frac{1}{|\mathcal{Q}_i|}\sum_{Q\in\mathcal{Q}_i}\frac{1}{n}\sum_{i=1}^n \hat{F}_{P_{iQ}|X_{iQ},Z_i}(1-p|B_i(q)-s,q,Z_i) \tag{E.6}$$

Distribution of the average *ex ante* costs Recall that:

$$\begin{aligned}
\bar{C}(q,a) &= \int_{\mathcal{C}^+}\left[1 - F_C(c;q,a)\right]dc - \int_{\mathcal{C}^-}\left[F_C(c;q,a)\right]dc \\
&= \int_{\mathcal{C}^+}\left[1 - m(c,q,a)\right]dc - \int_{\mathcal{C}^-}\left[m(c,q,a)\right]dc
\end{aligned}$$

Hence:

$$\Pr\left(\bar{C}(q,\alpha) \leq \bar{c}\right) = \int_0^1\int_{\mathcal{Z}} 1\left\{\int_{\mathcal{C}^+}\left[1-\widetilde{m}(c,q,z,\tau)\right]dc - \int_{\mathcal{C}^-}\left[\widetilde{m}(c,q,z,\tau)\right]dc \leq \bar{c}\right\} F_Z(dz)d\tau$$

Estimator:

$$\hat{P}\left(\bar{C}(q,a) \leq \bar{c}\right) = \frac{1}{|\mathcal{Q}_i|}\sum_{Q\in\mathcal{Q}_i}\frac{1}{n}\delta_\tau\sum_{i=1}^{n}\sum_{s=1}^{S_\tau} 1\left\{\hat{\hat{C}}(q,Z_i,\tau_s) \leq \bar{c}\right\} \tag{E.7}$$

$$\hat{\hat{C}}(q,Z_i,\tau) = \delta_c\sum_{s=1}^{S_c}\left(1\{y_s \geq 0\} - \widehat{\widehat{m}}(y_s,q,Z_i,\tau)\right) \tag{E.8}$$

$$\widehat{\widehat{m}}(y,q,Z_i,\tau) = \delta_p\sum_{s=1}^{S_p}\left(1 - 1\left\{\hat{F}_{P_{iQ}|X_{iQ},Z_i}(p_s|y,q,Z_i) \geq \tau\right\}\right) \tag{E.9}$$

F Comparison with a Linear Fixed-Effect Model

This section compares the results of the semiparametric model (SP) to the one from a linear fixed-effect model (FE) defined as follows:

$$\log\left(\frac{p_i(q)}{1-p_i(q)}\right) = \gamma \cdot B_{iQ} + \zeta q + \tau_i + u_{iq}, \;\; q = 0.01, 0.50, 0.90, Q_i \tag{F.1}$$

The dependent variable is the log odd of the intention to overstay. τ_i is an individual fixed-effect, u_{iq} an idiosyncratic disturbance.[16] The specification of Equation (F.1) is similar to Equation (9) in Wiswall and Zafar (2015), who introduce linear fixed-effects in a regression of the log odds (see Section 3 therein). In the model described in Section 3, Equation (3.4) yields Equation (F.1) under the following assumptions:

(i) $\nu^G_{i0} - \nu^E_{i0}$ follows an extreme-value type I distribution,

(ii) $\gamma = \mathbb{E}(\gamma_i|B_i(q), q)$ and $\zeta = \mathbb{E}\left(\zeta_{i1}|B_i(q), q\right)$,

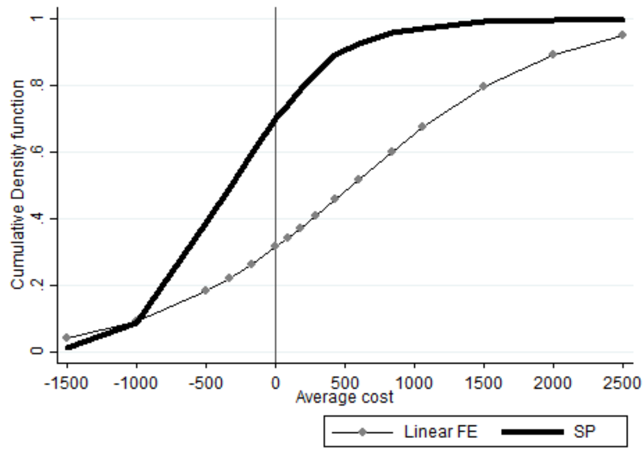
(iii) $\tau_i + u_{iq} = \zeta_{i0} - \mathbb{E}\left(\gamma_i \cdot B_i(q) + \zeta_{i1}\cdot q|B_i(q), q\right)$.

Wiswall and Zafar (2015) performs a Least-Absolute Deviation estimation under the assumption that the conditional median of u_{iq} equals 0. Note that this assumption is violated when q and u_{iq} are correlated, which is likely given expression (iii). In other words, the linear-fixed effect specification may not net out the individual utility component because it is not separable from q.

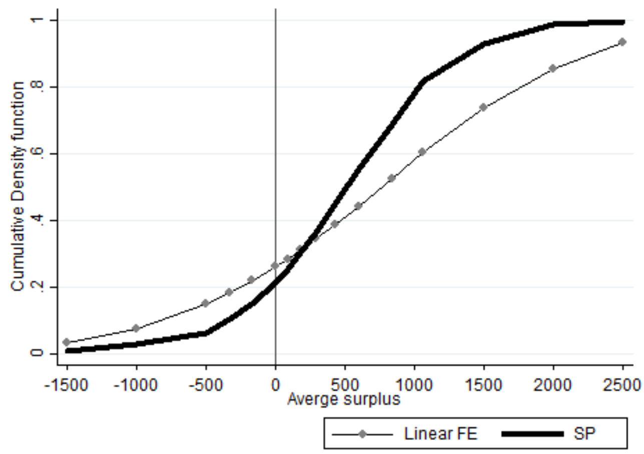
From Equation (F.1), the average *ex ante cost*, surplus and OV can be estimated as:

$$\begin{aligned}\hat{\hat{C}}_{iq} &= 1/\hat{\gamma}\cdot\left(\hat{\zeta}q + \hat{\tau}_i\right)\\ \hat{\hat{S}}_{iq} &= B_i(q) - 1/\hat{\gamma}\cdot\left(\hat{\zeta}q + \hat{\tau}_i\right)\\ \widehat{OV}_{iq} &= \hat{\hat{S}}_{iq} - \hat{\hat{S}}_{i0}\end{aligned}$$

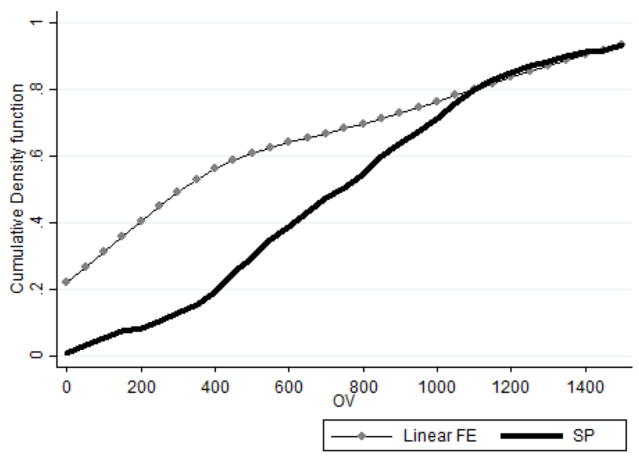
[16]The estimation results are qualitatively similar when q is interacted with an indicator variable for the city of residence.



(a) Cost



(b) Surplus



(c) OV

Figure F.1: Distribution of *ex ante* costs, surpluses and option-value: SP model vs. FE model

Figure F.1 presents the CDFs estimated using both procedures.[17] The linear FE estimation overestimates the average cost distribution, as the "linear FE" CDF dominates the "SP" CDF. This results in a distribution of surpluses with a fatter tail under the FE model, and an underestimation of the OV.

[17]The linear FE specification includes also an interaction term between q and an indicator for the city of residence.

G Additional Tables and Figures

		Germany	Berlin	Hamburg	Bavaria
2016	Open status	68%	71%	48%	67%
	Recognized	27%	24%	48%	26%
	Denied	5%	5%	4%	7%
2017	Open status	41%	42%	24%	41%
	Recognized	51%	52%	72%	51%
	Denied	8%	6%	5%	9%
2018	Open status	30%	29%	18%	30%
	Recognized	61%	62%	76%	61%
	Denied	9%	8%	7%	9%
2019	Open status	22%	20%	13%	21%
	Recognized	66%	68%	80%	69%
	Denied	12%	12%	7%	11%

Source: authors' calculation from DESTATIS

Table G.1: Distribution of status among Afghan migrants by German federal states and year

	Berlin	Hamburg	Munich	Total
Female	0.40	0.49	0.24	0.38
	(0.49)	(0.50)	(0.43)	(0.49)
Age	31.16	33.95	30.30	31.56
	(12.11)	(13.37)	(11.24)	(12.25)
Low-Skilled	0.69	0.57	0.63	0.65
	(0.46)	(0.50)	(0.48)	(0.48)
Years in DE	3.47	3.76	3.67	3.58
	(1.37)	(1.25)	(0.95)	(1.26)
Sampled from register	0.22	0.31	0.22	0.24
	(0.41)	(0.46)	(0.41)	(0.43)
Prev. occupied (Afg.)	0.54	0.50	0.50	0.52
	(0.50)	(0.50)	(0.50)	(0.50)
Secure Status	0.55	0.74	0.50	0.58
	(0.50)	(0.44)	(0.50)	(0.49)
Obtained educ. in DE	0.30	0.27	0.30	0.30
	(0.46)	(0.45)	(0.46)	(0.46)
No germ. class	0.16	0.15	0.28	0.19
	(0.37)	(0.36)	(0.45)	(0.39)
Germ. class (up to A2)	0.39	0.48	0.35	0.40
	(0.49)	(0.50)	(0.48)	(0.49)
Germ. class (B1 and more)	0.44	0.36	0.37	0.41
	(0.50)	(0.48)	(0.48)	(0.49)
Curr. Occupied	0.13	0.17	0.32	0.19
	(0.34)	(0.37)	(0.47)	(0.39)

Note: Mean values calculated on non-missing observations. Berlin N=534, Hamburg N=226, Munich N=264, Total N= 1,024. Standard deviation in parentheses. "Female" equals one if the respondent identifies as a female. "Low-skilled" equals one if the respondent has studied at most until lower secondary education. "Prev. occupied (Afg.)" refers to a previous occupation held in country of origin before migration. "Secure status" equals one if the respondent has received some form of temporary or permanent protection status. German class level B1 is the lower intermediate level from the Common European Framework of Reference for Languages (CEFR) standard, A2 the upper beginner level.

Table G.2: Sample characteristics by city

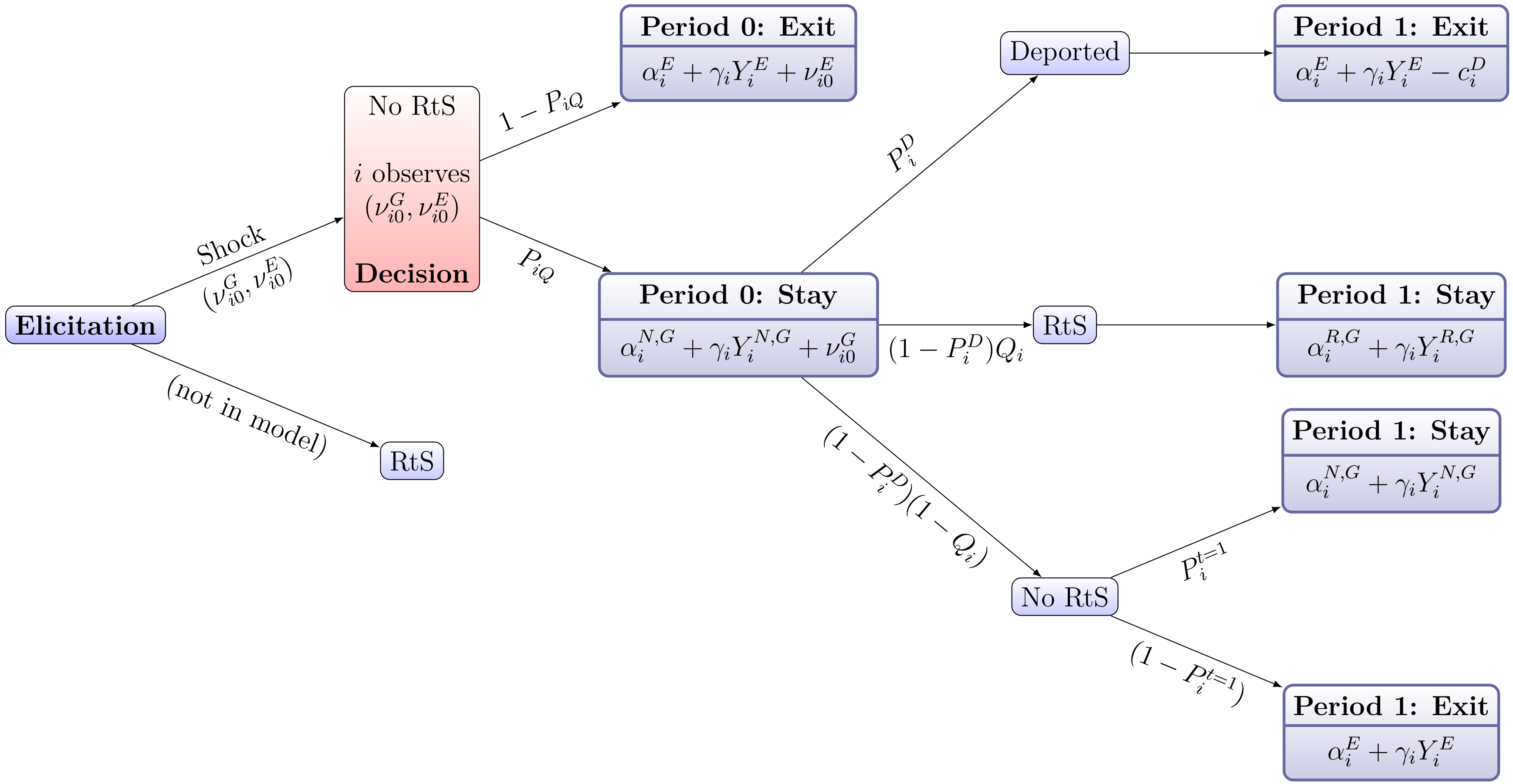
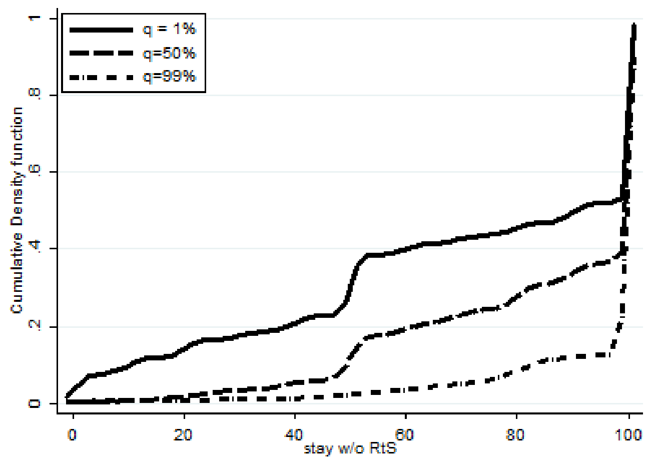
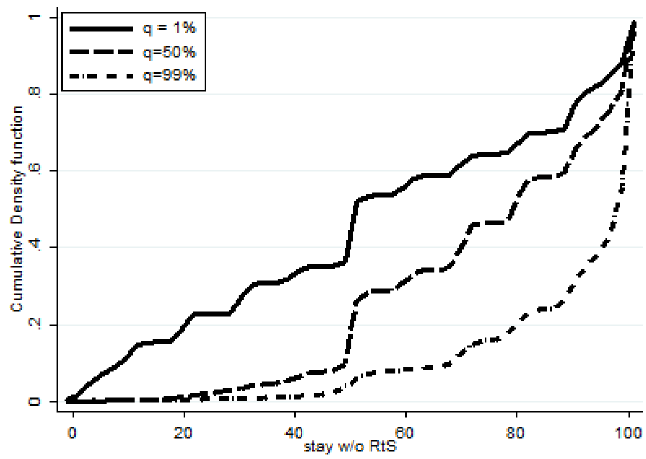


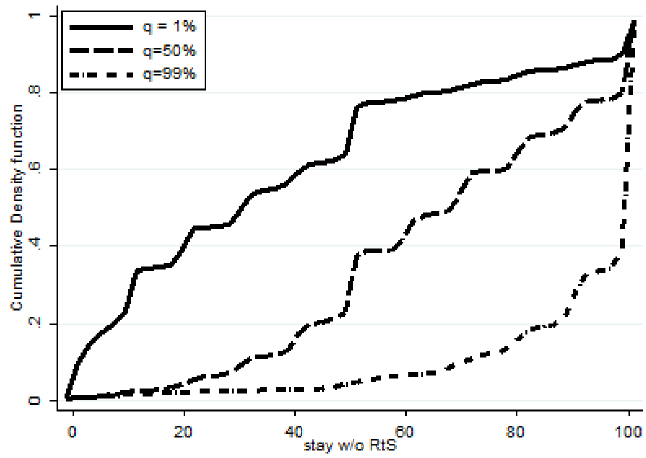
Figure G.1: Schematic representation of the simplified model of Section 3



(a) CDFs - Berlin



(b) CDFs - Hamburg



(c) CDFs - Munich

Figure G.2: Intention to stay w/o RtS by city for $q = 0.01, 0.50, 0.99$

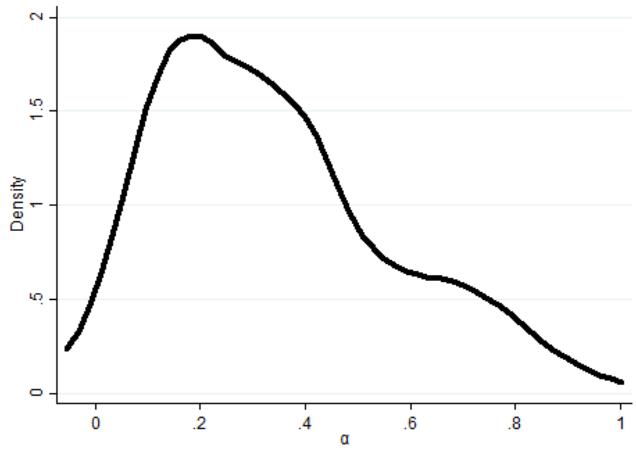


Figure G.3: Distribution of $\hat{\alpha}$

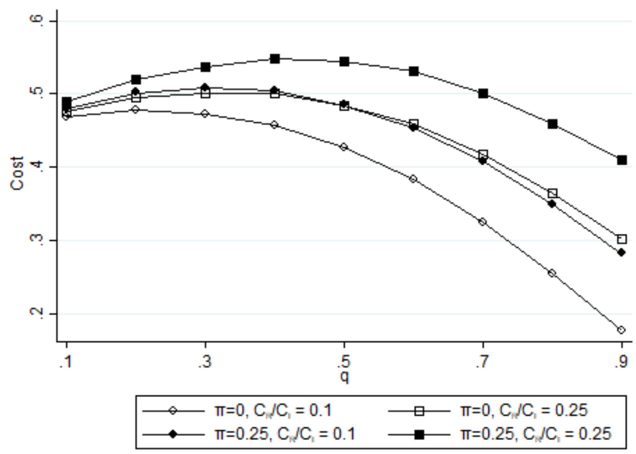


Figure G.4: Social Planner's cost function for different parameter values