

Explaining Differences Between the Expected and Actual Duration Until Return Migration: Economic Changes and Behavioral Factors

PRELIMINARY, please do not cite without permission

Gerard J. van den Berg*

Michèle A. Weynandt†

February 11, 2013

Abstract

This paper builds a duration model to get the realized return for non-returners. We then explore the difference between expectations and realizations in return migration. The analyses are based on the German Socio Economic Panel consisting of 25 years of information. The result lends support to the fact that people use simplifying heuristics when trying to forecast the future; their return intentions indicate bunching in heaps of 5 years (e.g. intend to return in 5, 10, 15 years). Along these lines we find that migrated individuals systematically underestimate the length of their stay in the receiving country. Furthermore we use behavioral factors to explain the difference between the intentions and the realized return. Here we find that the difference decreases the older one gets, but is larger the more disadvantaged one feels due to ones origin as an example. The robustness checks show that the results do not hinge on a single definition, or set of explaining variables. The consistency in the underestimation may have important policy and modeling implications.

JEL-Classification: C41, D03, D12, D84, F22

Keywords: Behavioral Economics, Duration Analysis, International Migration, Hedonic Forecasting, Utility Misprediction

*University of Mannheim, IFAU-Uppsala, VU University Amsterdam and IZA, Department of Economics; L 7, 3-5; 68131 Mannheim; Germany; [gerard\[at\]uni-mannheim\[dot\]de](mailto:gerard[at]uni-mannheim[dot]de)

†University of Mannheim, CDSE; L 7, 3-5; 68131 Mannheim; Germany; [michele\[dot\]weynandt\[at\]gess\[dot\]uni-mannheim\[dot\]de](mailto:michele[dot]weynandt[at]gess[dot]uni-mannheim[dot]de); tel. +49 621 181 3431. We greatly appreciate Annette Bergemann and David Card's help to get access to the data. We are grateful to Matthew Rabin, Stefano DellaVigna, Andrea Weber, Lena Janys, Johannes Schoch, Vera Molitor and seminar participants in the junior research dialogue at the University of Mannheim, SMYE 2012, SOEP 2012, for many helpful comments on an earlier version of this paper.

1 Introduction

This paper explores the fact that migrated individuals underestimate the length of their stay in the receiving country. “Hedonic forecasting” refers to the errors that individuals make in predicting changes in their tastes and feelings in the psychological literature. The reader is presented with evidence of a forecasting error and convincing statistics proving that it is not just simple noise. Loewenstein, O’Donoghue, and Rabin (2003) have defined the suggestion that people understand the qualitative nature of changes in their tastes, but the underestimation of the magnitude of these changes as projection bias.

Looking at return migration and the expectation to return, our prior is that people underestimate their attachment to the country of migration - when first moving away from home, one compares everything to home. Most of the time, the culture in the country of migration will be different, one will not know a lot of people and one may not even have family in the migrating country. All these things are examples of what one misses when first moving to a country. Furthermore as recently discussed in Card, Dustmann, and Preston (2012), prejudices from natives against migrants may hamper the adaptation and the process of feeling at home in Germany. Therefore when asked about whether or not one wants to return, most people say yes because they miss the culture, the food and so on¹.

Once one has fully arrived in the migrating country - Germany for the current analysis - one starts to meet new people, one gets to know people on the job - assuming that you have a job - and one starts to discover things about Germany that one may not have known in advance. This process of integrating and feeling at home in Germany is what is understood by net attachment in the following. When first coming to Germany, the net attachment is very low, even though one decided to migrate. The decision why people migrated in the first place underlies the current analysis and the focus lies on those migrants that are already in Germany.

The German Socio-Economic Panel (GSOEP) is used for the analysis as individuals provide information on their return intentions. Using a duration model an expression for the predicted return realization - an expected duration of the stay in the receiving country, Germany for the current analysis - is inferred. This predicted return will then be compared to the respondents intentions and will then be regressed on different sets of socio-economic variables, which allows for the identification² of the driving factors between return intentions and return realizations.

A first important finding, is that people’s intentions exert bunching which already points towards

¹Individuals that came to Germany due to a war or as refugees on the other hand may not want to return to their country ever. These individuals are of no worry for the current analysis, since they should predict that they want to stay in Germany forever.

²Please be aware that we are not claiming a causality of the results. We are only interested in the driving factors of the forecast error.

the fact that a simplifying heuristic may be at work. Taking a closer look at the difference between the intentions and the realizations, we see that the intentions lie constantly below the realization. Individuals considerably underestimate the duration of their stay. The average forecast error³ is therefore mostly negative but decreases the longer one stayed in Germany and the older one gets. Using pooled OLS, we are then also able to identify a few other factors that drive the difference between intentions and realizations. Being older than 60 years, reduces the difference considerably, while if an individual feels disadvantaged due to her origin, her forecast error increases. An individual, who is remitting over the course of her stay, is also underestimating the duration of her stay, while someone who has a high locus of control is better at predicting the duration of their stay.

The clear understanding of the difference between expectations and realizations in return migration is crucial for integration policies. If migrants consistently underestimate the duration of their stay, they may not put enough effort into their integration. Government intervention may help to improve the situation for migrants by emphasizing on integration as early as possible. It is important to understand these differences to avoid conflicts of integration between current inhabitants and migrants.

The setup of the paper is as follows; section 2 gives an overview of the relevant literature in return migration and ‘hedonic’ forecasting, projection bias. Section 3 presents the data and some preliminary results, while section 4 presents the model and the empirical specification. Section 5 presents the results and section 6 concludes.

2 Literature

The literature overview is split into two subsections, where first return migration is discussed, then the REH is explained with its relevant literature.

2.1 Return Migration

This subsection reviews a few groundbreaking papers in the field of return migration, which provide the underlying economic framework of the decision process; whether an individual should return or not. A first paper working out the details of return migration is the work of Borjas and Bratsberg (1996) who generalize the model of Borjas (1991) by allowing migrants to return. Borjas and Bratsberg (1996) mention two possible alternatives for return migration; one possibility is that return migration is part of the life-cycle and a second possibility is that the initial decision is based

³The difference between the intentions and the predicted return and forecasting error will be used interchangeably in the following since they refer to the same measure.

on erroneous information about economic opportunities in the receiving country, which then forces migrants to revise their information and return. Borjas and Bratsberg (1996) work focuses on the first possibility; the life-cycle argument.

Dustmann (2003b) complements Borjas and Bratsberg (1996) by adding two reasons for re-migration; either the returner has a relatively high preference for consumption at home or there is a higher purchasing power of the host country's currency in the sending country.

Likewise Dustmann (2003a) examines return motives of migrant parents and finds that parents who have a daughter are more likely to return to their home country than those that have a son. He explains his finding through the importance in cultural differences when raising a child. In other words, Dustmann (2003a) uses an altruistic model to show that "parental concerns about the child may lead to an increase or to a decrease in the tendency to return to the home country".

Dustmann and Weiss (2007) ream the above cases that return migration may occur because of a preference for home country consumption, a decision which would increase the migrants lifetime wealth. Along the lines of Borjas and Bratsberg (1996) life-cycle argument, Dustmann and Weiss (2007) claim that the benefits of migration decrease over the migration cycle, while costs are positive and may even increase. Dustmann, Bentolila, and Faini (1996) expand Borjas and Bratsberg (1996) life-cycle criteria by asserting that migrants may acquire skills in the receiving country that could be more valuable in their home country. As such the receiving country would only be a stop in their life-cycle. This reasoning goes along the lines of selective outmigration, where a recent example would be Van Hook and Zhang (2011) who find that emigration is positively associated with factors such as having a spouse in another country.

Another strand in the literature discusses the duration of stay and migratory frequency, usually illustrated by migration between Mexico and the United States (Hill (1987), Lindstrom (1996), Reyes (2001), Reyes (2004), Hill and Wong (2005), Durand, Kandel, Parrado, and Massey (1996)). Mexican migrants are frequent migrants, since they cross the border several times for a short period of time. They make about 4 or 5 trips and on average stay 6 months to a year per trip (Cornelius (1978), Jenkins (1977)).

The distinguishing feature of the current work is that it focuses on the underestimation of the trip duration.

The aforementioned literature discussed reasons for return migration, and as such constitutes the underlying component for the current work. Section 2.2 presents the concepts of 'hedonic' forecasting and projection bias.

2.2 Hedonic Forecasting and Projection Bias

A first and often cited example, in the hedonic forecasting and projection bias literature, is the work of Read and van Leeuwen (1998) regarding the prediction of hunger. They asked a group of

hungry and a group of satiated people what kind of snack - healthy or unhealthy - they wanted in a week at a time where both groups would be satiated (in the afternoon). Read and van Leeuwen (1998) found that the satiated group opted for the healthy snack while the hungry group preferred the unhealthy snack. Another paper on the same topic, Gilbert, Michael, Gill, and Wilson (2002) looked at people who were hungry and suggested that they acted as if their future taste for food would reflect such hunger. Nisbett and Kanouse (1968) suggested that shopping on an empty stomach may lead people to buy too much. Not just studies of hunger showed evidence of projection bias; Badger, Bickel, Giordano, Jacobs, Loewenstein, and March (2007) studied 13 long time adult heroin addicts who had been regularly receiving BUP⁴ and noticed that their expectations differed from the realized craving. Based on this evidence, Loewenstein, O'Donoghue, and Rabin (2003) formalized projection bias in predicting future utility.

It is well known that people adapt to changes, but the above cited literature presented evidence that people underestimate adaptation. Conlin, O'Donoghue, and Vogelsang (2007) clearly demonstrated how people exert projection bias by analyzing catalog orders. They were able to show that people were more likely to return winter clothes when the temperature on the receiving date climbed compared to the order date temperature. Gilbert, Pinel, Wilson, Blumberg, and Wheatley (1998) reported several instances of people underestimating adaptation to unfavorable events (which they labeled immune neglect). A very recent paper by Levy (2009) was able to pin down the projection bias in tobacco consumption. Furthermore Acland and Levy (2010) suggested that gym goers in an incentivized gym-use experiment do not appreciate the positive addiction of exercise regimes.

Stephens (2004) on the other hand examined the relationship between job loss expectations and realizations, and as such his focus is closer to the one considered in the current work. His work has two important outcomes; first he found that people's expectation were a good predictor of actual job loss. He found a positive correlation between the intention and the actual state, as such the expectation contained information that the econometrician could not infer from the demographics or other covariates. Second he discovered that workers in the HRS tended to overstate their job loss probability which is another important finding as one can see the connection to the underestimation of net attachment.

This paper contributes to the above mentioned literature by showing that people exert not just a prediction bias in food related issues, clothing or employment, but also in migration decisions. In addition, the goal of this work is to analyze people's ability to adapt their expectations over

⁴BUP stands for buprenorphine which is a drug that acts by relieving the symptoms of opiate withdrawal.

<http://www.employee-drug-testing-ace.com/employment-drug-screening-resources/employee-drug-testing-glossary/define-buprenorphine-bup>.

time and possibly show that their expectations converge to the truth in the long run. Levy (2009) and Acland and Levy (2010) look at habit formation over time and are able to show that people underestimate their addiction.

3 Data and Preliminary Results

Subsection 3.1 presents the Data, while subsection 3.2 provides evidence of projection bias in people's expectations.

3.1 Data

This paper uses the German Socio-Economic Panel⁵ (henceforth GSOEP) to analyze the difference between return expectations and return realizations of migrants to Germany. The GSOEP is a representative longitudinal survey of households and their members, whose aim is to collect representative micro-data on persons, households and families in order to measure stability and change in living conditions⁶. The GSOEP annually re-interviews households and their split-offs, usually in February and March. The current version has data from 1984 until 2010 which enables the duration analysis approach. The sample of the first wave (1984) includes about 1500 households with a foreign born head. Furthermore the GSOEP surveys the respondents intention, an important point in order to analyze the bias, by asking migrants about their desire to remain in Germany. First the respondent is asked whether she wants to return home, which can be answered by yes or no (stay in Germany forever). If she plans to return, there are two possible answers: "return within 12 months" or "return in a few years". If the plan is to return "in a few years" an intended amount of years that she plans to remain in Germany has to be provided to the interviewer.

The information about whether or not people return to their country of origin is provided by the GSOEP in the so called "address log" - where reasons for non-response are logged. The "address log" is recorded at the household level and has as possible options; "moved obtained address", "address of the household not found", "address unknown", "moved out of Germany" or "died". "Moved out of Germany" is used to code the migrant's return.

Using the return status we are able to infer the expected return through duration analysis, the predicted expectations will be compared to the given intentions⁷. GDP is used as a proxy for the life conditions in the home country and as a proxy of the possible wage in the sending country which is necessary to infer the predicted return. The GDP levels for the different countries are from

⁵To get a more thorough overview of the data, we refer you to Wagner, Frick, and Schupp (2007).

⁶Haisken-DeNew and Frick (2005).

⁷We refer you the section 4 for further details on the duration model for the expected return.

Angus Maddison⁸ but are only available until 2008, which forces the drop of the year 2009 and leaves 25 years for the analysis (1984-2008). The Maddison data was chosen because it incorporates most countries of origin for the migrants in the current sample. Furthermore the GDP levels are in 1990 International Geary-Khamis (GK)\$.

3.2 Presence of a Bias

To illustrate the actual returns Table 1 presents the number of returns across the years. People that have not returned until 2009 are coded as non returners for Table 1 and all upcoming results. From a duration analysis point of view these observations are right censored. As can be seen in Table 1 overall close to 23 % return over the course of 25 years, from 1985 until 2009, while on the annual level about 1% of individuals return⁹.

[TABLE 1 ABOUT HERE]

A comparison between the actual and the intended return provides evidence that people's expectations differ from their actions. Evidence that people may exhibit projection bias in forecasting their future is presented in Table 2¹⁰. About 70 percent of those expressing the intention to return to their home country, over the course of 25 years never do¹¹. As mentioned above when evaluating Table 2 keep in mind that some people may have been wrongly coded as non returners. They can still return but it cannot be observed due to right censoring. A further thing to note, is that it is impossible to capture short term migration lasting no longer than one year. The GSOEP surveys people annually, thereby not allowing the account of people that migrate and return within a year¹².

[TABLE 2 ABOUT HERE]

⁸[http://www.ggd.net/MADDISON/Historical Statistics/horizontal-file 02-2010.xls](http://www.ggd.net/MADDISON/Historical%20Statistics/horizontal-file%2002-2010.xls).

⁹Table 1 should look similar to Table 1 in Dustmann (2003a), as you can see by comparing my table with his, which is reproduced in the Appendix A.1, my numbers are smaller than his. The code was double checked, in order to make sure that individuals not households were considered, nevertheless our results differ. In the Appendix we discuss possible explanations for these differences.

¹⁰In the Appendix (Table A.2) the comparison between the intended and the actual return from 1984-1997 is provided in order to make it possible to compare these results to Dustmann (2003a), but again the numbers differ.

¹¹As can be seen by comparing Table 2 with Table A.3 the overall numbers do not change much when the time horizon is enlarged by 12 years, from 1984-1997 to 1984-2009.

¹²These individuals do not play an important role for the analysis of the underestimation of the trip duration.

A valid concern in assessing the above numbers is that individuals do not report the truth to the interviewer when asked about their desire to return. Some people may lie about their planned duration in Germany because their current visa only allows them to stay for a limited amount of time. Since the GSOEP provides information on a migrant's residence status, which is either unlimited or limited, Table 3 presents the comparison between the desire to return and the residence permit question. About 70 percent of those that have a limited residence permit in Germany reply that they want to remain in Germany. As a consequence one cannot argue that people tend to lie due to their residence permit. As it may be easy to get the residential permit prolonged people respond truthfully when asked about their intentions¹³. Be aware that in Table 3 the information that is available across all years from 1984 until 2009 is used, while Table 2 only considers those people that are present in 1984. Unfortunately it is not possible to present a table with those individuals present in 1984, since for everyone of them the residence status is missing - an unfortunate side effect of survey data¹⁴.

[TABLE 3 ABOUT HERE]

As we have seen up to this point, there is evidence of a bias between people's expectation and their final actions in the case of return migration. Table 4 takes a closer look at the socio-economic differences between movers and stayers.

[TABLE 4 ABOUT HERE]

There seem to be socio-economic differences between movers and stayers, a finding which goes along with the findings of e.g. Van Hook and Zhang (2011). Leavers and stayers seem to differ in certain socio-economic characteristics, e.g. marital status, employment, which points toward the selection of return migrants¹⁵.

¹³Note that in Table 3 there are three different possible answers for the desire to return home, while in Table 2 the intention to return home was coded as a yes or no. If people answered that they want to stay in Germany, their intentions to stay was coded as a yes, while if people answered that they either plan to return within 12 months or after 1 year, their intentions to stay were coded as a no.

¹⁴Tables A.4 and A.5 in the appendix include the same baseline year, a group of people for whom the residence permit status is known and for whom the intentions are known. 1996 is the first year this happens which shortens the time horizon notably.

¹⁵Table A.6 in the appendix splits the "stayers" into attriters and those individuals that we observe until 2008 and have not returned yet. Attrition is often discussed when a panel data set is used, but hopefully does not play a significant role here.

[FIGURE 1 ABOUT HERE]

[FIGURE 2 ABOUT HERE]

Figure 1 and 2 contrast the descriptive statistics of the four possible groups. In our sample we have leavers who never say that they want to remain forever, leavers who at some point say that they want to remain forever and non leavers who either say that they want to remain forever at some point or never say that they want to stay forever. Figure 1a) shows the number of observations for the different groups. The group that at some point said that they wanted to remain forever and have not yet left constitute the largest group. Figure 1b) then takes a closer look at the gender composition of the different groups. There seem to be no significant differences in gender between the different groups. Figure 1c) then looks at whether children are present. Here we see that those individuals that have children are more likely to be in the group that says at some point that they want to remain forever and have not left so far. Figure 1d) then looks at whether there are significant gender differences for the children between the different groups. We thought at first that there may be a difference, since some parents may want their girls to grow up in their home culture, while for their boys, they would prefer the German environment since it may constitute a better working environment. But as panel d) shows, there seem to be no such differences. Figure 2 then continues to contrast different characteristics, but there seem to be no relevant differences between the four groups. Panel a) takes a closer look at the unemployment versus employment rates, panel b) looks at differences between marital status, panel c) contrasts the languages spoken at home, while panel d) graphs the different educational levels of the individuals. As already mentioned, there seem to be no significant differences between the four groups in terms of these characteristics. So none of these characteristics should drive the differences in between the intentions and the expectations in the following.

The next section provides the reader with the methodology used to infer the actual return based on the current information available to the individual.

4 Model

Let T be the duration until the return and let $\theta(t|x(t), x_0)$ be the hazard rate, which can be interpreted as the return rate or the return probability. Mathematically it can be represented as:

$$\theta(t, x(t), x_0) = \lim_{dt \rightarrow 0} \frac{P(t \leq T < t + dt | T \geq t, x(t), x_0)}{dt} \quad (1)$$

t presents time since entry, $x(t)$ are time varying covariates, such as the current employment status, the current family income, and x_0 are time invariant covariates, such as the age at migration,

gender, education, country of origin.

The amount of money that migrants will earn in their home country and how the purchasing powers differ between the migrants country of origin and Germany builds the framework for the analysis between expectations and realizations. Information about what migrants wages would be in their home country is not available and GDP is used to infer how big the differences are between Germany and the sending country. Since the focus of the paper is to explain differences between return intentions and return realizations, we need an expression for the return realization which will be inferred through duration analysis. This analysis is said to be reduced form and we need to think about possible factors that migrants consider when forming expectations.

GDP is a good indicator to compare countries and as mentioned in the literature review the decision to return may be a part of the life-cycle, or the sending country may have caught up to the receiving country in terms of GDP. Comparing the GDP's of Germany and that of the sending countries, we know that either this did not happen, e.g. for countries such as Turkey, or Germany was just as good in terms of GDP as the sending country, e.g. France. In other words, a change in the arguments of the utility function changes the utility level. This can be modeled with the help of the duration analysis. To do so, first assume that the migrants to Germany are a homogeneous group, an assumption which may be relaxed in future work.

As emphasized above, the decision to return relies on the economic model which builds the framework for the hazard rate. As an example, for an individual to take the decision to move in 2005 it is needed that the expected present value of earnings proxied by GDP in the home country minus the moving costs are larger than the expected present value of earnings proxied by GDP in Germany. This formulation of the decision to move has been introduced by Sjaastad (1962). More formally, if one decides to move in 2005,

$$\sum_{t=2005}^d \frac{1}{(1+r)^t} (E[U(X_T(t))] - E[U(X_G(t))]) > c + \epsilon \quad (2)$$

needs to hold. Where $X(t)$, are covariates that we control for, such as GDP, age at migration, marital status, family location c represents the cost of moving, d is the expected year of death, r is the interest rate and ϵ is an error term. The subscript G stands for Germany, and the subscript T stands for Turkey¹⁶. This can be rewritten in terms of probabilities, such that:

$$\begin{aligned} &P(\text{move in 2005 from Germany to Turkey}) \quad (3) \\ &= P\left(\epsilon < \sum_{t=2005}^d \frac{1}{(1+r)^t} (E[U(X_T(t))] - E[U(X_G(t))]) - c\right). \end{aligned}$$

¹⁶Turkey was chosen as an example, since as can be seen in Table A.7 most migrants in the sample are from Turkey.

Which can also be rewritten in terms of the hazard rate in 2005, such that:

$$\begin{aligned}
P \left(\epsilon < \sum_{t=2005}^d \frac{1}{(1+r)^t} (E[U(X_T(t))] - E[U(X_G(t))]) - c \right) \\
= \Phi \left(\frac{\sum_{t=2005}^d \frac{1}{(1+r)^t} (E[U(X_T(t))] - E[U(X_G(t))]) - c}{\sigma_\epsilon} \right)
\end{aligned} \tag{4}$$

Equation (6) is the expression of the hazard for 2005 and can easily be rewritten to get an expression for the hazard rate for each year.

Since we are ultimately interested in the expected duration of a stay, the duration framework allows us to write:

$$\begin{aligned}
y(0) &= E(T|x_0, \text{expectations of future path of } x(t)) \\
&= \int_0^\infty \left[\exp \left(- \int_0^\infty \theta(u|x(u), x_0) du \right) \right] dz
\end{aligned} \tag{5}$$

in a continuous time framework. This equation can be rewritten for $y(t)$ where t can take any integer value in $[0, T]$ which means that we end up with possible $y(t)$, $y(t-1)$, ..., $y(0)$. This expression allows the individual to adapt her expectations. In other words, $y(0)$ may be different than $y(1)$ because people update the future path of $x(t)$. The model's predicted expectations will be compared to the respondents indicated intentions to see what drives the difference and whether people learn; are their predictions eventually converging to the "truth"?

Empirical Specification

Since the data at hand are of the discrete time format, the expected duration until the return will be based on the assumption of a third order polynomial duration dependence¹⁷ combined with a complementary log log model. Then the full model specification would be (assuming time invariant covariates):

$$\text{cloglog}[h(t, X)] = z_1 t + z_2 t^2 + z_3 t^3 + \beta X \tag{6}$$

where X represents socio-economic characteristics¹⁸. In other words, the hazard can be rewritten as:

$$h(t, X) = 1 - \exp[-\exp(z_1 t + z_2 t^2 + z_3 t^3 + \beta X)] \tag{7}$$

¹⁷The third order polynomial is our preferred specification of the duration dependence, see Table 5 and the results section for more details.

¹⁸We control for sex, age at migration, difference in GDP between Germany and the source country, marital status, whether or not the individual attended school in Germany, whether or not the individual has family at home and whether or not the individual's spouse is at home. Furthermore we control for the country of origin.

where z_1, z_2, z_3 are estimated together with the intercept and the slope parameters within the vector β . Survival up to the end of the j^{th} interval (or completion of the j^{th} cycle) is given by:

$$S(j) = S_j = \prod_{k=1}^j (1 - h_k) \quad (8)$$

where h_k is the cloglog function of characteristics.

For each individual, we want to calculate the expected duration of the stay at the moment of the interview. Thus even if the interview happens when a person has already spent 10 years in Germany, we want to calculate the expected duration of the stay from that point onwards. Therefore we consider the year of the interview as $t = 0$. Consider now the case where people form their expectations based on the current GDP only, and all other variables included in the model so far do not vary with time or only vary once - marital status, employed, family at home, spouse at home. Age at migration and attended school in Germany are really time invariant covariates. Hence the predicted return in the discrete time framework is given by,

$$E[T] = \sum_{k=1}^K S(k) = S(1) + S(2) + S(3) + \dots + S(K) \quad (9)$$

where K is the maximum survival time¹⁹. The predicted return can be rewritten as:

$$E[T] = (1 - h_1) + (1 - h_1)(1 - h_2) + \dots + (1 - h_1)(1 - h_2)(1 - h_3) \dots (1 - h_K) \quad (10)$$

where h_x represent the hazard at time x ²⁰. In the following the predicted return will be denoted by $E[T]$ while the intended duration will be denoted by $\tilde{E}[T]$. The next subsection discusses the results for this model and explains the sample selection criteria.

5 Results

As shortly mentioned in the data section, we consider only the migrants that are already in Germany and present in the GSOEP. Furthermore we only consider adults, who are older than 18 years in order to only include those individuals that take the return decision themselves. As the use of the GDP Data from Angus Maddison forced the drop of the year 2009, we are left with 25

¹⁹In the empirical part we assume that the maximum survival time equals the expected lifetime duration, approximated by 100 - current age

²⁰As an example:

$$\begin{aligned} h_1(t, X) &= 1 - \exp[-\exp(z_1 t + z_2 t^2 + z_3 t^3 + \beta X)] \\ h_2(t + 1, X) &= 1 - \exp[-\exp(z_1(t + 1) + z_2(t + 1)^2 + z_3(t + 1)^3 + \beta X)] \\ h_3(t + 2, X) &= 1 - \exp[-\exp(z_1(t + 2) + z_2(t + 2)^2 + z_3(t + 2)^3 + \beta X)] \end{aligned}$$

years for the analysis (1984-2008) and 3152 individuals, where 574 durations until re-migration are not right censored.

Tables 5 and 6 show the results of the complementary log log model, and logit model which are the underlying models for the predicted return. These specifications allow the construction of the predicted return as stated in the methodology section. The estimates are shown to provide evidence that all the coefficients point in the right direction. As an example, being employed makes you less likely to return, while having your spouse in the your home country makes you more likely to return. Males also seem to be less likely to return than females. Compared to singles every other marital status type is less likely to return. Whether the logit model or the complementary clog log specifications are used, does not change these effects. Most of the coefficients are significant at the 1% level.

Furthermore Table 5 as well as Table 6 test which duration specification may be the best. In both Tables, Column (1) includes year dummies, in order to give a fully nonparametric specification of the duration dependence, while Column (5) includes Time Interval Dummies, allowing for a piecewise constant specification of the duration dependence. We also tried $\ln(t)$ ²¹ as well as a fifth order polynomial in time and a third order polynomial in time.

Our preferred specification is the third order polynomial, which also seems to fit the pattern that at the beginning one may be more likely to return, while then the likelihood to return decreases until one reaches the retirement age, where the likelihood increases again. These specifications, as explained in the methodology section, allow us to “extract” the hazard rate which allow the construction of the predicted return. All predicted returns analyzed below are based on the complementary log log model with a third order polynomial in time to model the duration dependence.

[TABLE 5 ABOUT HERE]

[TABLE 6 ABOUT HERE]

Before analyzing the differences between the intentions and the predicted return, let us look at the individuals intentions and what may be driving factors of changes in these intentions. Figure 3 plots the Intended Duration of Stay, in panel a) we imputed the intended duration for those who wanted to stay forever as 100 – their current age, while in panel b) we only take a look at those that actually tell us how long they plan on staying. In both panels we can see that the individuals show bunching behavior around 5, 10, 15, 20 years. This bunching may already point towards a simplifying heuristic, when individuals form their intentions.

²¹Which is the discrete-time analogue of the continuous time Weibull model.

[Figure 3 ABOUT HERE]

Tables 7 and 8 then look at the driving factors behind the changes in peoples return intentions. We take the first difference in their intentions - as an example we compute $\tilde{E}[2006] - \tilde{E}[2005]$ - and regress these changes in their intentions on the changes in their socio-economic changes; e.g. $\text{employed}_{2006} - \text{employed}_{2005}$. All regressions include individual fixed effects and the standard errors are clustered at the individual level. What seems to be a driving factor in these adjustments is whether there is a change in your life satisfaction, meaning that if you are more satisfied in one year than in the following²², it influences your intention to return. This finding is as expected, since an increase in life satisfaction may also reduce your psychic costs that occur from migration. Other variables that seem to have significant effect on these changes are the Attended School in Germany variable and the differences in GDP variable. All the above mentioned variables coefficients are significant at the 1 % level.

[TABLE 7 ABOUT HERE]

Table 8 then takes a closer look at some of the behavioral factors contained in the GSOEP and how they influence the changes in the individuals intentions. Unfortunately the number of observations decreases substantially depending on which variables are included. Data on control over life is only available in years 1994-1996, 1999 and 2005, data on remitting is only available in years 1984-1993 and 1995, while data on risk preferences is only available in years 2006-2009²³. Column (1) is the baseline specification without behavioral factors, in order to make it easier for the reader to see what happens to the sample size when the other variables are included. Another unfortunate side-effect of the small sample size is that none of the behavioral coefficients are significant, which does not give us much room for argumentation.

[TABLE 8 ABOUT HERE]

Last but not least, let us move to the forecast errors. Figure 4 plots the difference between the intentions and the predicted return. Again panel a) plots the difference for the whole sample where for those that intended to stay forever we imputed their maximal survival time as 100 – their current age. Panel b) plots the difference for the reduced sample where we leave those out that intend to stay forever. A quick glance at panel a) gives us hope, that there seem to be many

²²Happiness variable.

²³To get a little more observations, we fill the variables forward using stata's `stfill` command. This means that the variable takes on the last value until a change in the variable happens. This assumption should not change any of the results we find.

people predicting the duration of their stay correctly, but when we take those out that intend to stay forever²⁴(panel b)) practically no remaining individual has a correct prediction. Figure 4b) shows that individuals overestimate the return to their home country, equivalently stated, underestimate their time spent in Germany. When looking at the difference, the intended return is constantly below the actual predicted return, which makes the difference negative. This is an important finding and may point toward overconfidence; a topic very nicely introduced in Kahneman (2011). In the current work, overconfidence would have to go along with net attachment in the sense that individuals are overconfident about the fact that they will be true to their family (to their “roots”) and want to return home, and thus underestimate their attachment to Germany. More importantly though the finding of the overestimation of the probabilities to return is along the lines of Rabin (2002). He models the belief in the “Law of Small Numbers”, where people exaggerate the degree to which small samples resemble the population from which they are drawn.

[FIGURE 4 ABOUT HERE]

A further thing we looked at was whether “narrow framing” plays a role. We use the term “narrow framing” to define the fact that people may only look at shorter time intervals than their whole lives. Trying to predict what will happen in 40 years from now is hard, and therefore it may also be hard to predict what one will do in 40 years concerning the return migration. Therefore we take a look at people’s predictions if we make a cut off at e.g. 10 years and everyone that states that they want to remain longer than 10 years, we re-code them as only wanting to stay 10 years²⁵. The top panel of figure 5 looks at the intended duration of stay, once we restrict the window for predictions to 10 years. Panel a) includes those that intend to stay forever (coded as intending to stay for 10 more years), while panel b) excludes these individuals. The lower panel of figure 5 then takes a look at the difference between the narrowly framed intention and the predicted return of this model. In panel c), close to 70% of the individuals now predict their stay correctly, but again taking those out that intend to stay forever, only 35% of the individuals seem to correctly predict the length of their stay within this framework. Nevertheless it is important to notice that in Figure 4 panel b) nearly nobody predicted the length of their stay correctly. These results point towards the fact that not just overconfidence may play a role here, but also the forecasting ability of the individuals. It is easier to give a response to what you may be doing in a year or two than give a response to the question about when you may want to return. For the rest of the paper we will go back to the initial specification of the predicted return, where the sum is taken until the expected

²⁴Which are those that we imposed on how long they remain in Germany.

²⁵We played around with these numbers, and 10 just gave the best predictions, which is why we stick to that number.

survival time, which is approximated by $100 - \text{the current age of the individual}$.

[FIGURE 5 ABOUT HERE]

Another important question when looking at these graphs is whether people actually learn something. In other words, is the difference between their intentions and the predicted realization approaching zero the longer they stay in Germany? Figure 6 tries to take a closer look at this learning problem by looking at the changes over time spent in Germany. Looking at the different panels, it seems like people do not learn to predict their preferences more accurately the longer they are in Germany. The distribution shifts a little bit closer to zero which may be due to the fact that the population gets older.

[FIGURE 6 ABOUT HERE]

Figure 7 then plots the average forecast error (equivalent to the average difference) over different time specifications. The average forecast error is the important measure here, as the REH analyzes exactly this average. Panel a) plots the average forecast error over time spent in Germany, and what was not obvious before now seems to become relevant. It seems that the longer people actually are in Germany, the more accurate they get on average. The largest error that they make is when they have spent 20 years in Germany, while their error is practically zero once they have spent 60 years in Germany. This could go along with the fact that having spent 20 years in a country you may still believe that you eventually return, but the older you get, the better you are at comparing your actual chance of leaving and so you seem to be more accurate with your forecast.

Panel b) then helps us explain at what age you seem to get better at predicting your utility or your future choice variables. Toward this end, there is a clear direction; the older you get, the better you get at predicting your remaining duration. This finding is also not a too surprising as the older one is, the shorter the remaining horizon gets, and therefore one may also be better at predicting the duration of the stay.

Panel c) plots the average forecast error over the different Life Satisfaction possibilities, where 0 stands for not satisfied, while a 10 ranks you at very satisfied. We decided to also show this graph, since happiness seemed to explain some of the changes in the intended returns, but how satisfied you are with your current life does not seem to have a different effect on the average forecast error.

Panel d) then plots the average forecast error over time. The increasing slope does not make too much sense to us, except that the individuals that we consider in the sample may get older by these years and as already stated above then get better at forecasting their own preferences. We nevertheless thought it was useful to include, in order to show that there do not seem to be relevant

macro shocks that may drive our results.

[FIGURE 7 ABOUT HERE]

Tables 9,10, 11 and 12 then finally take a closer look at the differences between the intentions and the predicted return. In table 9, columns (2), (4) and (5) include individual fixed effects, where the standard errors are also clustered at the individual level. Nearly every coefficient is significant at the 1% level, and the more we control for, some of these significance levels drop, but nevertheless stay significant at the 10% level. The OLS results, are shown just for comparison, since we are in a panel data set up, we need to include individual fixed effects and also cluster the standard errors²⁶. As an example, one can see the effect when taking a closer look at attended school in Germany. The coefficient changes sign and magnitude as soon as we include individual fixed effects and clustering. A further thing to note, is that the coefficient estimate on time spent in Germany are all significant and point into the right direction. The longer one has been in Germany, the smaller the difference between the intentions and the predicted return. The coefficient on the above 60 dummy is also highly significant, showing that it is very important to control for this hump at the retirement age. The coefficient on the above 60 dummy is positive, but since the difference is always negative, this means that the difference actually decreases as one is above 60. The coefficient on the disadvantage due to origin variable is also significant at the 1% level, and is negative. This implies, again as the difference is on average negative, that those individuals that feel a disadvantage underestimating the duration of their stay by more than those that do not feel disadvantaged. Having attended school in Germany actually is one of the surprising coefficients since it increases the difference when we focus on the specifications that include individual fixed effects in table 9 columns(3), (5) and (6).

Table 10 and Table 11 are robustness checks of the results from Table 9. Table 10 is a first test on the sign and the magnitude of the results of Table 9, as we exclude those observations where individuals indicate that they intend to stay forever. For Table 11 we split the sample randomly in half, where for one half of the sample the hazard model was estimated while for the other half, the average forecast error was predicted using the results from the hazard model.

Table 12 then again includes the behavioral factors where we take a look at the whole sample again, but as before we lose out on power as we lose observations. Here the interesting behavioral results come from whether or not a person has remitted and again the Locus of Control variable. If an individual has been paying remittances, she is underestimating the time she is going to spent in Germany. The coefficient is negative, but again as the difference is always negative, this means

²⁶The identification with the use of individual fixed effects is driven by variations across time by each individual. Since many of the variables included in the regression may be time invariant, we included the yearly OLS results in the Appendix in Tables A.8 and A.9.

that in absolute terms the difference becomes larger. On the other hand, if an individual has control over his or her life, they will be better at giving an estimate of their duration of stay. The coefficient is positive, indicating that the difference becomes smaller. The coefficients on the risk aversion are unfortunately not significant.

[TABLE 9 ABOUT HERE]

[TABLE 10 ABOUT HERE]

[TABLE 11 ABOUT HERE]

[TABLE 12 ABOUT HERE]

6 Conclusion

The current article showed evidence of a difference between expectations and realizations of the duration of the stay in the host country. Unfortunately we were not able to show whether there is projection bias, due to data restrictions, but we showed that predictions get better the longer one stayed in Germany. The main rationale behind this finding, in our opinion, is that the individuals time horizon that they have left to live, shortens every year that they have spent in Germany and therefore their prediction gets better and more accurate. This also goes along the findings of Smith, Taylor, and Sloan (2001), who using the Health and Retirement Survey (HRS), found that longevity expectations are consistently linked to subsequent observed mortality. The participants of the HRS have reached a certain age, and therefore their evidence coincides with ours²⁷. Another interpretation of the results leads towards Kahneman (2011) description of what you see is all there is (WYSIATI). As shortly mentioned when we presented the results of people's intentions, there seems to be bunching at 5, 10, 15 years, which points toward a simplifying heuristic at work. WYSIATI goes into the same direction. When you ask people about returning to their home country, things they like about their culture / home country become more salient. This in turn may also make their wish to return more salient and thereby bias the given answer.

In the introduction we mentioned that the findings would be relevant for government action. As it is not clear what really drives these differences, we need to be careful when giving policy advice. In future research we want to figure out where policy interventions would help, and whether the

²⁷Another example that uses the HRS is Sergeant, Ekerdt, and Chapin (2010) who analyzed retirement migration and found that expectations to move correctly predicted the move into a community, but did not predict the move into nursing facilities.

intentions that people provide really coincide with their future actions taken. As an example, if an individual thinks that she will return in less than five years, she may not start to integrate properly. As it turns out, this individual will stay longer than she at first thought. The time that the individual spent thinking that she may return quicker could have therefore been used more efficiently, as an example for Germany, the individual could have started to learn German.

To conclude, this article presents relevant information about the fact that migrants underestimate their stay in the country of origin, but there also seems to be a learning effect. The longer they are in the host country, the older they become and the better their forecasts become.

References

- ACLAND, D., AND M. LEVY (2010): "Habit Formation and Naivete, and Projection Bias in Gym Attendance," *mimeo*.
- BADGER, G., W. BICKEL, L. GIORDANO, E. JACOBS, G. LOEWENSTEIN, AND L. MARCH (2007): "Altered states: The impact of immediate craving on the valuation of current and future opioids," *Journal of Health Economics*, September 2007, 26(5), 865–876.
- BORJAS, G. J. (1991): *Immigration and Self-Selection in: Immigration, trade and the labor market*. University of Chicago Press, Chicago, IL.
- BORJAS, G. J., AND B. BRATSBERG (1996): "Who leaves? The outmigration of the foreign-born," *Review of Economics and Statistics*, 78(1), 165–176.
- CARD, D., C. DUSTMANN, AND I. PRESTON (2012): "Immigration, wages and compositional amenities," *Journal of the European Economic Association*, 10 (1), 78–119.
- CONLIN, M., T. O'DONOGHUE, AND T. J. VOGELSANG (2007): "Projection Bias in Predicting Catalog Orders," *American Economic Review*, 97(4), 1217–1249.
- CORNELIUS, W. A. (1978): *Mexican Migration to the United States: Causes, Consequences, and U.S. Responses*. Migration and Development Study Group Monograph C/76-12. Cambridge: Massachusetts Institute of Technology, Center for International Studies.
- DURAND, J., W. KANDEL, E. A. PARRADO, AND D. S. MASSEY (1996): "International Migration and Development in Mexican Communities," *Demography*, 33(2), 249–264.
- DUSTMANN, C. (2003a): "Children and Return Migration," *Journal of Population Economics*, 16, 815–830.
- (2003b): "Return Migration, wage differentials, and the optimal migration duration," *European Economic Review*, 47, 353–369.
- DUSTMANN, C., S. BENTOLILA, AND R. FAINI (1996): "Return Migration: The European Experience," *Economic Policy*, 11(22), 213–250.
- DUSTMANN, C., AND Y. WEISS (2007): "Return Migration: Theory and Empirical Evidence from the UK," *British Journal of Industrial Relations*, 45:2, 236–256.
- GILBERT, D., T. MICHAEL, J. GILL, AND T. D. WILSON (2002): "The Future is Now: Temporal Correction in Affective Forecasting," *Organizational Behavior and Human Decision Processes*, LXXXVIII, 430–444.

- GILBERT, D. T., E. C. PINEL, T. D. WILSON, S. J. BLUMBERG, AND T. P. WHEATLEY (1998): "Immune Neglect: A source of Durability Bias in Affective For," *Journal of Personality and Social Psychology*, LXXXV, 617–638.
- HAIKEN-DENEW, J. P., AND J. R. FRICK (2005): "DTC Desktop Companion to the German Socio-Economic Panel (SOEP)," Version 8.0.
- HILL, J. K. (1987): "Immigrant Decisions Concerning Duration of Stay and Migratory Frequency," *Journal of Development Economics*, 25, 221–234.
- HILL, K., AND R. WONG (2005): "Mexico-US Migration: Views from Both Sides of th Border," *Population and Development Review*, 31(1), 1–18.
- JENKINS, C. J. (1977): "Push/Pull in Recent Mexican Migration to the U.S.," *International Migration Review*, 11(2), 178–189.
- KAHNEMAN, D. (2011): *Thinking, Fast and Slow*. Farrar, Straus and Giroux.
- LEVY, M. (2009): "An Empirical Analysis of Biases in Cigarette Addiction," *mimeo*.
- LINDSTROM, D. P. (1996): "Economic Opportunity in Mexico and Return Migration from the United States," *Demography*, 33(3), 357–374.
- LOEWENSTEIN, G., T. O'DONOGHUE, AND M. RABIN (2003): "Projection Bias in Predicting Future Utility," *Quarterly Journal of Economics*, 118(4), 1209–1248.
- NISBETT, R. E., AND D. E. KANOUSE (1968): "Obesity, Hunger, and Supermarket Shopping Behavior," *Proceedings of the Annual Convention if the American Psychological Association*, III, 683–684.
- RABIN, M. (2002): "Inference By Believers in the Law of Small Numbers," *Quarterly Journal of Economics*, 117(3), 775–816.
- READ, D., AND B. VAN LEEUWEN (1998): "Predicting Hunger: The Effects of Appetite and Delay on Choice," *Orga*, LXXXVI, 189–205.
- REYES, B. I. (2001): "Immigrant Trip Duration: The Case of Immigrants from Western Mexico," *International Migration Review*, 35(4), 1185–1204.
- (2004): "Changes in Trip Duration for Mexican Immigrants to the United States," *Population Research and Policy Review*, 23, 235–257.
- SERGEANT, J. F., D. J. EKERDT, AND R. K. CHAPIN (2010): "Older Adults' Expectations to Move: Do They Predict Actual Community-Based or Nursing Facility Moves Within 2 Years?," *Journal of Aging and Health*, 22(7), 1029–1053.

- SJAASTAD, L. A. (1962): “The Costs and Returns of Human Migration,” *Journal of Political Economy*, 70(5), 80–93.
- SMITH, K. V., D. H. TAYLOR, AND F. A. SLOAN (2001): “Longevity Expectations and Death: Can People Predict Their Own Demise?,” *The American Economic Review*, 91(4), 1126–1134.
- STEPHENS, M. J. (2004): “Job Loss Expectations, Realizations and Household Consumption Behavior,” *The Review of Economics and Statistics*, 86(1), 253–269.
- VAN HOOK, J., AND W. ZHANG (2011): “Who Stays? Who Goes? Selective Emigration Among the Foreign-Born,” *Population Research and Policy Review*, 30, 1–24.
- WAGNER, G., J. FRICK, AND J. SCHUPP (2007): *The German Socio-Economic Panel Study (SOEP) - Scope, Evolution and Enhancements*. Schmollers Jahrbuch 127(1)139-169.

Table 1: **Return Frequency**

Year	85	86	87	88	89	90	91	92	93	94	95	96	97
Return	117	33	37	53	33	28	17	23	24	29	27	41	23
Pct	3.70	1.04	1.17	1.68	1.04	0.89	0.54	0.73	0.76	0.92	0.85	1.30	0.73
Year	98	99	00	01	02	03	04	05	06	07	08	09	Total
Return	39	27	24	23	15	11	23	14	16	13	17	14	721
Pct	1.23	0.85	0.76	0.73	0.47	0.35	0.73	0.44	0.51	0.41	0.54	0.44	22.82

Source: GSOEP, Own Calculations.

Table 2: **Intentions and Realization 1984 - 2009**

Return between 84 and 09			
Intended Return (84)	No	Yes	Total
No	682	82	764
Column Percentage	30.00	16.05	27.44
Row Percentage	89.27	10.73	
Yes	1591	429	2020
Column Percentage	70.00	83.95	72.56
Row Percentage	78.76	21.24	
Total	2273	511	2784

This Table only presents statistics for people present in 1984.

Source: SOEP, own calculations.

Table 3: **Desire to Return versus Residence Status**

Desire to Return	Residence Status		
	Unlimited	Limited	Total
Within 12 Months	22	24	46
(Percentage)	0.79	1.48	1.04
After One Year	766	444	1210
Percentage	27.47	27.44	27.46
Stay in Germany Forever	2000	1150	3150
Percentage	71.74	71.08	71.49
Total	2788	1618	4406

Source: GSOEP, own calculations.

Table 4: Socioeconomic Differences

Variable	Stayers			Leavers			t-stat
	Mean	SD	N	Mean	SD	N	
Male	0.50	0.50	3891	0.44	0.50	574	(-2.56)*
Age at Migration	30.04	10.66	3891	30.79	9.34	574	(1.59)
$\ln(\text{GDP}_G) - \ln(\text{GDP}_H)$	1.69	1.10	3838	1.69	1.06	568	(0.05)
Married	0.65	0.48	3564	0.38	0.49	471	(-11.50)***
Married living separated	0.02	0.15	3564	0.02	0.14	471	(-0.43)
Divorced	0.05	0.22	3564	0.01	0.12	471	(-3.54)***
Widowed	0.05	0.22	3564	0.03	0.16	471	(-2.56)*
Employed	0.52	0.50	3890	0.44	0.50	574	(-3.65)***
Family at Home	0.19	0.39	3876	0.07	0.25	569	(-7.32)***
Spouse at Home	0.02	0.13	3891	0.08	0.27	574	(9.25)***
Attended School in Germany	0.03	0.17	3832	0.02	0.16	566	(-0.85)

Source: GSOEP, own calculations.

Note: The t-statistics test for the significance of the difference between leavers and stayers. For each individual the last point in time where information is provided in the data set is taken to get the different means.

Table 5: Complementary Log-log model

	(1)	(2)	(3)	(4)	(5)
leave					
Male	-0.146 (0.107)	-0.148 (0.107)	-0.143 (0.107)	-0.143 (0.107)	-0.145 (0.107)
Age at Migration	-0.00390 (0.00692)	-0.00510 (0.00652)	-0.00332 (0.00691)	-0.00335 (0.00693)	-0.00687 (0.00672)
ln(GDP _G)-ln(GDP _H)	0.00845 (0.0547)	0.0102 (0.0546)	0.0131 (0.0547)	0.0130 (0.0547)	0.00609 (0.0545)
Married	-0.646*** (0.144)	-0.590*** (0.121)	-0.643*** (0.142)	-0.642*** (0.143)	-0.552*** (0.132)
Married living separated	-0.565 (0.389)	-0.464 (0.377)	-0.544 (0.387)	-0.543 (0.387)	-0.441 (0.381)
Divorced	-1.097*** (0.405)	-1.039*** (0.394)	-1.098*** (0.404)	-1.096*** (0.405)	-0.965** (0.398)
Widowed	-1.115*** (0.389)	-1.064*** (0.374)	-1.136*** (0.388)	-1.135*** (0.389)	-1.002*** (0.381)
Employed	-0.691*** (0.157)	-0.694*** (0.157)	-0.674*** (0.157)	-0.674*** (0.157)	-0.681*** (0.157)
Family at Home	0.0289 (0.198)	-0.00108 (0.197)	-0.00941 (0.198)	-0.00973 (0.198)	-0.0343 (0.197)
Spouse at Home	1.095*** (0.190)	1.127*** (0.189)	1.120*** (0.189)	1.120*** (0.189)	1.103*** (0.189)
Attended School in Germany	-0.142 (0.287)	-0.169 (0.286)	-0.150 (0.286)	-0.150 (0.286)	-0.155 (0.286)
GDP _G Growth	0.0234 (0.0284)	0.0232 (0.0282)	0.0259 (0.0283)	0.0259 (0.0283)	0.0202 (0.0284)
GDP _H Growth	0.0242** (0.0115)	0.0274** (0.0119)	0.0262** (0.0118)	0.0262** (0.0118)	0.0255** (0.0116)
GDP _G Growth Imy	-0.00894** (0.00350)	-0.00764*** (0.00247)	-0.00849** (0.00338)	-0.00852** (0.00342)	-0.00641** (0.00287)
GDP _H Growth Imy	0.00410*** (0.000845)	0.00406*** (0.000797)	0.00396*** (0.000828)	0.00397*** (0.000831)	0.00420*** (0.000828)
Income	0.0000748 (0.0000593)	0.0000768 (0.0000588)	0.0000753 (0.0000591)	0.0000753 (0.0000591)	0.0000723 (0.0000592)
Aged 60 or older	1.035*** (0.173)	1.064*** (0.166)	1.022*** (0.172)	1.023*** (0.173)	1.105*** (0.169)
ln(t)		-0.484*** (0.129)			
Time in Germany			-0.192*** (0.0480)	-0.196** (0.0865)	
Time in Germany ²			0.00742*** (0.00240)	0.00774 (0.00623)	
Time in Germany ³			-0.0000922*** (0.0000339)	-0.000100 (0.000145)	
Time in Germany ⁵				1.31e-09 (2.36e-08)	
Constant	-20.01 (723.6)	-3.630*** (0.432)	-3.526*** (0.402)	-3.512*** (0.471)	
Year Dummies	Yes	No	No	No	No
Country Region	Yes	Yes	Yes	Yes	Yes
Time Interval Dummies	No	No	No	No	Yes
Observations	32200	32200	32200	32200	32200

Source: GSOEP, own calculations.

Note: *, **, *** indicates significance at the 10%, 5%, and 1% level, respectively. Standard errors in parentheses. The dependent variable is the dummy variable whether a person leaves or not.

Table 6: Logit model

	(1)	(2)	(3)	(4)	(5)
leave					
Male	-0.147 (0.109)	-0.148 (0.109)	-0.143 (0.109)	-0.143 (0.109)	-0.145 (0.109)
Age at Migration	-0.00379 (0.00708)	-0.00503 (0.00664)	-0.00323 (0.00704)	-0.00328 (0.00707)	-0.00684 (0.00687)
ln(GDP _G)-ln(GDP _H)	0.00820 (0.0559)	0.0109 (0.0555)	0.0133 (0.0557)	0.0132 (0.0557)	0.00580 (0.0556)
Married	-0.660*** (0.146)	-0.598*** (0.123)	-0.653*** (0.144)	-0.652*** (0.145)	-0.559*** (0.134)
Married living separated	-0.575 (0.402)	-0.461 (0.389)	-0.541 (0.399)	-0.539 (0.399)	-0.444 (0.393)
Divorced	-1.115*** (0.411)	-1.055*** (0.398)	-1.114*** (0.408)	-1.112*** (0.409)	-0.978** (0.402)
Widowed	-1.142*** (0.395)	-1.081*** (0.379)	-1.155*** (0.393)	-1.153*** (0.393)	-1.019*** (0.386)
Employed	-0.704*** (0.159)	-0.705*** (0.158)	-0.685*** (0.159)	-0.685*** (0.159)	-0.694*** (0.159)
Family at Home	0.0263 (0.201)	-0.00243 (0.200)	-0.0106 (0.201)	-0.0111 (0.201)	-0.0365 (0.200)
Spouse at Home	1.144*** (0.199)	1.162*** (0.197)	1.156*** (0.197)	1.155*** (0.197)	1.144*** (0.197)
Attended School in Germany	-0.142 (0.292)	-0.172 (0.290)	-0.154 (0.291)	-0.154 (0.291)	-0.160 (0.290)
GDP _G Growth	0.0237 (0.0293)	0.0233 (0.0289)	0.0261 (0.0291)	0.0261 (0.0291)	0.0200 (0.0292)
GDP _H Growth	0.0248** (0.0120)	0.0280** (0.0123)	0.0268** (0.0122)	0.0268** (0.0122)	0.0260** (0.0120)
GDP _G Growth Imy	-0.00918** (0.00358)	-0.00776*** (0.00251)	-0.00862** (0.00344)	-0.00866** (0.00347)	-0.00647** (0.00291)
GDP _H Growth Imy	0.00416*** (0.000866)	0.00414*** (0.000815)	0.00403*** (0.000847)	0.00404*** (0.000850)	0.00429*** (0.000847)
Income	0.0000760 (0.0000599)	0.0000780 (0.0000594)	0.0000765 (0.0000597)	0.0000764 (0.0000597)	0.0000739 (0.0000597)
Aged 60 or older	1.053*** (0.176)	1.080*** (0.169)	1.038*** (0.175)	1.040*** (0.175)	1.124*** (0.172)
ln(t)		-0.495*** (0.132)			
Time in Germany			-0.196*** (0.0492)	-0.203** (0.0885)	
Time in Germany ²			0.00761*** (0.00245)	0.00814 (0.00636)	
Time in Germany ³			-0.0000946*** (0.0000347)	-0.000108 (0.000147)	
Time in Germany ⁵				2.16e-09 (2.39e-08)	
Constant	-19.40 (679.1)	-3.607*** (0.441)	-3.498*** (0.412)	-3.475*** (0.483)	
Year Dummies	Yes	No	No	No	No
Country Region	Yes	Yes	Yes	Yes	Yes
Time Interval Dummies	No	No	No	No	Yes
Observations	32200	32200	32200	32200	32200

Source: GSOEP, own calculations.

Note: *, **, *** indicates significance at the 10%, 5%, and 1% level, respectively. Standard errors in parentheses. The dependent variable is the dummy variable whether a person leaves or not.

Table 7: Difference in Expectations

	(1)	(2)	(3)	(4)	(5)	(6)
ln(GDP _G)-ln(GDP _H) FD	-1.35*	-1.35*	-1.48*	-2.10**	-1.21	-1.88**
	(0.82)	(0.82)	(0.82)	(0.82)	(0.83)	(0.83)
Employed FD	0.22*	0.22*	0.21	0.25*	0.20	0.19
	(0.13)	(0.13)	(0.13)	(0.13)	(0.13)	(0.13)
Income FD	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Family at Home FD	-0.02	-0.09	-0.11	-0.06	-0.05	-0.14
	(0.49)	(0.50)	(0.50)	(0.49)	(0.46)	(0.51)
Spouse at Home FD	-0.29	-0.29	-0.29	-0.31	-0.40	-0.43
	(0.41)	(0.41)	(0.41)	(0.46)	(0.44)	(0.44)
Attended School in Germany FD	-0.12	-0.12	-0.54**	-0.28*	0.22	-0.55**
	(0.25)	(0.25)	(0.26)	(0.16)	(0.24)	(0.22)
Happiness FD	0.10***	0.10***	0.10***	0.09***	0.09***	0.10***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Have Children FD	0.42	0.42	0.22			-0.27
	(0.81)	(0.81)	(0.81)			(0.61)
Death of Mother FD		1.52				
		(1.15)				
Death of Father FD		-1.20				
		(0.86)				
Aged 60 or older			0.38***			0.49***
			(0.11)			(0.12)
Time in Germany			-0.04***			-0.06***
			(0.01)			(0.01)
Married				-0.45***		
				(0.10)		
Married living separated				-0.52		
				(0.37)		
Divorced				-0.24		
				(0.40)		
Widowed				-0.43*		
				(0.26)		
Finished Higher Education					-0.59	
					(0.48)	
Finished School					0.79	
					(0.78)	
vocational					-1.14	
					(0.74)	
Married FD						-0.15
						(0.25)
Married living separated FD						-0.07
						(0.72)
Divorced FD						-0.04
						(0.82)
Widowed FD						-0.06
						(0.29)
Constant	-0.13***	-0.13***	0.72***	0.00	-0.02	1.04***
	(0.01)	(0.01)	(0.16)	(0.03)	(0.06)	(0.18)
Number of Clusters	1858	1858	1858	1692	1783	1593
Observations	11219	11219	11219	10502	10867	9555
R ²	0.10	0.10	0.10	0.10	0.10	0.12

Source: GSOEP, own calculations.

Note: *, **, *** indicates significance at the 10%, 5%, and 1% level, respectively. Clustered standard errors in parentheses. The dependent variable is the change of the expected duration of stay. All regressions include individual fixed effects.

Table 8: Difference in Expectations Behavioral Factors

	(1)	(2)	(3)	(4)	(5)
ln(GDP _G)-ln(GDP _H) FD	-1.88** (0.83)	0.11 (1.73)	-1.46* (0.83)	5.03 (6.50)	58.51 (43.15)
Employed FD	0.19 (0.13)	0.01 (0.20)	0.21 (0.14)	0.07 (0.55)	-0.29 (0.41)
Income FD	-0.00* (0.00)	-0.00 (0.00)	-0.00 (0.00)	0.00** (0.00)	0.00** (0.00)
Spouse at Home FD	-0.43 (0.44)		-0.29 (0.41)		
Family at Home FD	-0.13 (0.50)				
Attended School in Germany FD	-0.62*** (0.17)				
Happiness FD	0.10*** (0.03)	-0.00 (0.07)	0.10*** (0.03)	-0.12 (0.23)	0.04 (0.14)
Aged 60 or older	0.49*** (0.12)	0.18 (0.21)	0.39*** (0.11)	0.51 (1.08)	0.25 (1.10)
Time in Germany	-0.06*** (0.01)	-0.01 (0.02)	-0.04*** (0.01)	0.05 (0.20)	-0.01 (0.24)
Married FD	-0.15 (0.25)	0.00 (0.31)		0.59 (1.46)	-0.31 (2.50)
Married living separated FD	-0.07 (0.72)	0.00 (0.79)		1.09 (1.72)	0.41 (2.56)
Divorced FD	-0.03 (0.82)	0.01 (0.93)		-0.29 (1.53)	-0.77 (2.30)
Widowed FD	-0.06 (0.29)	-0.31 (0.38)			
Control Over Life		0.18 (0.36)			
Remitting			-0.15 (0.14)		
Medium low risktaker				-0.15 (0.55)	
Medium high risktaker				-0.24 (0.82)	
High risktaker				-0.94 (0.70)	
Risktaker FD					-0.14 (0.44)
Constant	1.03*** (0.18)	-0.29 (0.66)	0.81*** (0.18)	-1.50 (5.30)	1.06 (6.78)
Number of Clusters	1593	723	1713	339	299
Observations	9555	3068	10798	901	700
R ²	0.12	0.12	0.09	0.12	0.34

Source: GSOEP, own calculations.

Note: *, **, *** indicates significance at the 10%, 5%, and 1% level, respectively. Clustered standard errors in parentheses. The dependent variable is the change of the expected duration of stay. All regressions include individual fixed effects. Data on Control over life is only available in years 1994-1996, 1999 and 2005. Data on remitting is only available in years 1984-1992, 1993 and 1995. Data on risk preferences is only available in years 2006-2009.

Table 9: Difference between the Intentions and the predicted Return

	(1)	(2)	(3)	(4)	(5)	(6)
Male	-0.92*** (0.30)	-0.14 (0.32)		-0.44 (0.32)		
Age at Migration	0.25*** (0.02)	0.21*** (0.02)		0.14*** (0.02)		
ln(GDP _G)-ln(GDP _H)	-1.31 (2.62)	-17.97*** (3.69)	-4.75 (3.63)	-6.18 (23.26)	-4.90 (3.69)	-5.04 (3.68)
ln(GDP(t-1) _G)-ln(GDP(t-1) _H)	19.30*** (4.29)	18.34*** (5.67)	13.23*** (4.30)	13.55 (23.68)	13.84*** (4.54)	13.94*** (4.54)
ln(GDP(t-2) _G)-ln(GDP(t-2) _H)	-16.71*** (2.66)	1.28 (3.58)	-6.25* (3.27)	-8.03 (6.02)	-6.37* (3.45)	-6.51* (3.45)
Married	13.56*** (0.31)	-4.43*** (0.80)	-2.45 (2.15)	-4.34*** (0.82)	-1.41 (2.34)	-1.36 (2.33)
Married living separated	13.71*** (1.14)	-3.77*** (1.29)	-1.99 (2.46)	-4.29*** (1.30)	-0.92 (2.71)	-0.77 (2.72)
Divorced	16.29*** (0.75)	-1.88* (1.00)	-4.09* (2.13)	-2.64*** (1.01)	-2.70 (2.45)	-2.57 (2.44)
Widowed	12.42*** (0.85)	-4.66*** (1.10)	-3.83* (2.30)	-4.90*** (1.13)	-2.33 (2.49)	-2.12 (2.49)
Employed	-0.48* (0.28)	-0.60* (0.33)	-0.82* (0.47)	-1.92*** (0.34)	-0.92* (0.49)	-0.88* (0.51)
Family at Home	7.46*** (0.40)	3.46*** (0.37)	7.29 (6.43)	2.77*** (0.38)	5.97 (5.85)	6.00 (5.83)
Spouse at Home	1.46* (0.85)	-3.65*** (1.48)		-1.24 (1.53)		
Attended School in Germany	3.53*** (0.68)	1.28 (0.83)	-25.28*** (0.99)	0.82 (0.83)	-25.17*** (1.06)	-25.35*** (1.09)
Time in Germany	-0.88*** (0.13)	0.44*** (0.15)	-0.67* (0.34)	-0.11 (0.17)	-0.64* (0.36)	-0.63* (0.36)
Time in Germany ²	0.03*** (0.01)	-0.05*** (0.01)	0.03** (0.02)	-0.01 (0.01)	0.03* (0.02)	0.03* (0.02)
Time in Germany ³	-0.00*** (0.00)	0.00*** (0.00)	-0.00* (0.00)	0.00*** (0.00)	-0.00 (0.00)	-0.00 (0.00)
Children?	-0.98*** (0.32)	1.52*** (0.40)	-3.08 (2.49)	1.24*** (0.40)	-3.41 (2.40)	-3.54 (2.40)
Aged 60 or older	4.07*** (0.47)	2.46*** (0.54)	3.76*** (0.69)	3.19*** (0.55)	3.55*** (0.73)	3.48*** (0.73)
Writing German?		3.27*** (0.45)	0.45 (0.78)	-0.07 (0.51)	0.61 (0.86)	0.63 (0.86)
Speaking German?		3.07*** (1.06)	-0.42 (1.88)	1.39 (1.32)	-0.62 (2.59)	-0.62 (2.60)
Disadvantage due to origin?				-1.42*** (0.30)	-0.78** (0.38)	-0.75** (0.37)
Language Newspaper German?				5.78*** (0.41)	0.93 (0.69)	0.93 (0.69)
Income						-0.00 (0.00)
Happiness						0.21** (0.09)
Constant	-26.87*** (1.09)	-19.66*** (1.65)	-5.46 (7.57)	-16.94*** (3.45)	-9.90 (7.89)	-11.14 (7.93)
Country Region	No	No	No	Yes	No	No
Bundesland FE	No	No	No	Yes	Yes	Yes
R ²	0.20	0.11	0.68	0.18	0.68	0.68
Number of Clusters			2075		1950	1950
Observations	26603	13258	13258	12336	12336	12336

Source: GSOEP, own calculations.

Note: ***,** indicates significance at the 10%, 5%, and 1% level, respectively. Standard errors in parentheses. The dependent variable is the difference between the intended return and the predicted realization. The columns that include the number of clusters, include individual fixed effects and those standard errors are clustered.

Table 10: Difference between the Intentions and the predicted Return without those that intend to stay forever

	(1)	(2)	(3)	(4)	(5)
Male	-0.56** (0.27)		-0.53* (0.28)		
Age at Migration	0.69*** (0.02)		0.65*** (0.02)		
ln(GDP _G)-ln(GDP _H)	-5.81* (3.04)	-7.70*** (2.23)	-9.36*** (3.07)	-7.98*** (2.28)	-7.94*** (2.27)
ln(GDP(t-1) _G)-ln(GDP(t-1) _H)	7.12 (4.37)	4.26** (2.05)	6.74 (4.36)	4.93** (2.06)	4.96** (2.06)
ln(GDP(t-2) _G)-ln(GDP(t-2) _H)	-0.94 (3.05)	2.91 (3.03)	2.65 (3.16)	1.54 (2.58)	1.54 (2.58)
Married	-6.18*** (0.63)	-6.79 (4.72)	-4.93*** (0.67)	-7.97* (4.68)	-8.01* (4.64)
Married living separated	-3.24*** (1.06)	-7.22* (4.25)	-2.91*** (1.08)	-8.28* (4.30)	-8.31* (4.27)
Divorced	-7.60*** (0.89)	-9.18** (4.46)	-5.79*** (0.91)	-10.75** (4.63)	-10.88** (4.63)
Widowed	-6.51*** (0.94)	-8.09* (4.53)	-6.88*** (1.00)	-9.50** (4.73)	-9.66** (4.72)
Employed	-0.03 (0.27)	-1.40*** (0.31)	-0.80*** (0.28)	-1.54*** (0.32)	-1.71*** (0.40)
Family at Home	0.62 (0.45)	-0.47 (1.26)	1.07** (0.51)	-0.74 (1.35)	-0.80 (1.40)
Spouse at Home	2.83*** (0.85)		4.09*** (0.89)		
Attended School in Germany	-1.91*** (0.72)	-2.40*** (0.55)	-1.77** (0.70)	-2.44*** (0.56)	-2.64*** (0.60)
Time in Germany	1.25*** (0.16)	0.98** (0.40)	1.39*** (0.16)	1.19*** (0.43)	1.17*** (0.43)
Time in Germany ²	-0.02** (0.01)	-0.01 (0.02)	-0.02*** (0.01)	-0.02 (0.02)	-0.02 (0.02)
Time in Germany ³	0.00** (0.00)	0.00 (0.00)	0.00** (0.00)	0.00 (0.00)	0.00 (0.00)
Children?	-0.41 (0.32)	-3.60 (2.78)	-0.24 (0.33)	-3.65 (2.81)	-3.53 (2.80)
Aged 60 or older	2.87*** (0.40)	3.78*** (0.56)	3.42*** (0.42)	3.70*** (0.61)	3.76*** (0.62)
Writing German?	1.48*** (0.31)	-0.40 (0.34)	0.60* (0.33)	-0.39 (0.36)	-0.39 (0.35)
Speaking German?	0.19 (0.72)	-0.33 (0.64)	0.53 (0.86)	-0.46 (0.81)	-0.50 (0.82)
Disadvantage due to origin?			-0.59** (0.25)	-0.06 (0.28)	-0.10 (0.29)
Language Newspaper German?			1.36*** (0.28)	0.38 (0.43)	0.37 (0.42)
Income					0.00 (0.00)
Happiness					-0.16** (0.08)
Constant	-74.04*** (1.45)	-43.61*** (9.20)	-77.16*** (1.86)	-42.43*** (9.87)	-41.14*** (9.81)
Country Region	No	No	Yes	No	No
Bundesland FE	No	No	Yes	Yes	Yes
R ²	0.69	0.91	0.72	0.91	0.91
Number of Clusters		760		705	705
Observations	3133	3133	2883	2883	2883

Source: GSOEP, own calculations.

Note: *, **, *** indicates significance at the 10%, 5%, and 1% level, respectively. Standard errors in parentheses. The dependent variable is the difference between the intended return and the predicted realization without those that intend to stay forever. The columns that include the number of clusters, include individual fixed effects and those standard errors are clustered.

Table 11: Difference between the Intentions and the predicted Return RO-BUST

	(1)	(2)	(3)	(4)	(5)
Male	-1.07*** (0.41)		-1.34*** (0.42)		
Age at Migration	0.14*** (0.03)		0.09*** (0.03)		
ln(GDP _G)-ln(GDP _H)	-7.17 (4.89)	10.14** (4.81)	-0.18 (4.93)	10.05** (4.76)	10.04** (4.75)
ln(GDP(t-1) _G)-ln(GDP(t-1) _H)	8.96 (7.61)	4.79 (5.40)	5.64 (7.50)	4.67 (5.41)	4.55 (5.44)
ln(GDP(t-2) _G)-ln(GDP(t-2) _H)	-1.58 (4.47)	-10.07*** (3.86)	-4.93 (4.44)	-9.84** (3.86)	-9.94** (3.86)
Married	-5.19*** (1.12)	1.98 (4.62)	-5.07*** (1.13)	2.49 (4.48)	2.32 (4.49)
Married living separated	-5.74*** (1.66)	-1.67 (4.71)	-5.18*** (1.65)	-1.23 (4.59)	-1.29 (4.63)
Divorced	-2.75** (1.38)	0.22 (4.43)	-3.04** (1.38)	0.71 (4.28)	0.67 (4.31)
Widowed	-6.91*** (1.44)	-1.35 (4.39)	-6.57*** (1.45)	-0.89 (4.26)	-0.84 (4.29)
Employed	-1.87*** (0.44)	-1.44** (0.61)	-2.38*** (0.44)	-1.42** (0.60)	-0.95 (0.62)
Family at Home	3.42*** (0.49)	16.96 (12.91)	2.78*** (0.51)	16.03 (12.32)	15.72 (12.39)
Spouse at Home	-2.45 (1.94)		-2.02 (1.92)		
Attended School in Germany	-0.14 (1.14)	-26.46*** (1.30)	0.83 (1.14)	-26.33*** (1.28)	-26.84*** (1.32)
Children?	3.85*** (0.52)	-0.48 (3.81)	3.46*** (0.52)	-1.33 (3.65)	-1.52 (3.63)
Aged 60 or older	4.08*** (0.70)	3.44*** (0.92)	4.45*** (0.71)	3.49*** (0.92)	3.32*** (0.93)
Time in Germany	0.60*** (0.21)	-0.73 (0.51)	0.40* (0.21)	-0.74 (0.49)	-0.73 (0.49)
Time in Germany ²	-0.05*** (0.01)	0.04** (0.02)	-0.04*** (0.01)	0.05** (0.02)	0.04** (0.02)
Time in Germany ³	0.00*** (0.00)	-0.00* (0.00)	0.00*** (0.00)	-0.00* (0.00)	-0.00* (0.00)
Disadvantage due to origin?	-1.92*** (0.39)	-1.52*** (0.51)	-1.73*** (0.39)	-1.46*** (0.51)	-1.42*** (0.51)
Writing German?	-1.23* (0.65)	0.20 (1.12)	-1.36** (0.65)	0.28 (1.09)	0.32 (1.08)
Speaking German?	-1.36 (1.78)	-0.70 (3.66)	-2.05 (1.76)	-0.55 (3.69)	-0.48 (3.70)
Language Newspaper German?	7.80*** (0.50)	2.08** (0.91)	6.11*** (0.55)	2.07** (0.91)	2.03** (0.92)
Income					-0.00** (0.00)
Happiness					0.18 (0.12)
Constant	-60.06*** (17.72)	38.12 (50.03)	-41.82** (18.04)	39.46 (50.47)	39.41 (50.19)
Country Region	No	No	Yes	No	No
Bundesland FE	No	No	Yes	Yes	Yes
R ²	0.14	0.67	0.17	0.67	0.67
Number of Clusters		1083		1083	1083
Observations	7129	7129	7098	7098	7098

Source: GSOEP, own calculations.

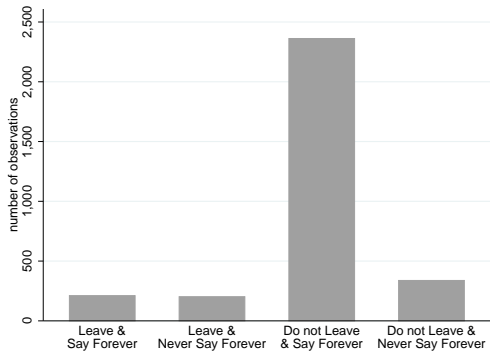
Note: *, **, *** indicates significance at the 10%, 5%, and 1% level, respectively. Standard errors in parentheses. The dependent variable is the difference between the intended return and the predicted realization. The columns that include the number of clusters, include individual fixed effects and those standard errors are clustered. Furthermore we have done random sampling to get half of the sample to deduce the coefficients and then imputed for the other half the predicted realization.

Table 12: Difference between the Intentions and the predicted Return Behavioral Factors

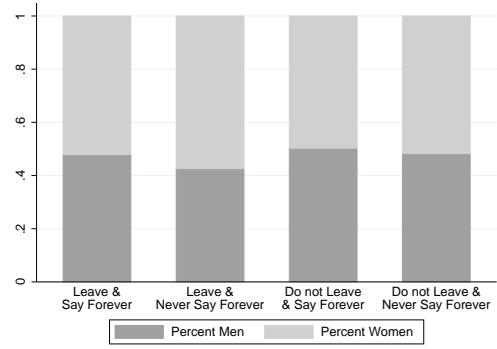
	(1)	(2)	(3)	(4)	(5)	(6)
$\ln(\text{GDP}_G) - \ln(\text{GDP}_H)$	-5.04 (3.68)	-6.94 (4.57)	-6.78** (2.91)	9.50*** (3.12)	4.67 (12.25)	4.17 (12.25)
$\ln(\text{GDP}(t-1)_G) - \ln(\text{GDP}(t-1)_H)$	13.94*** (4.54)	17.71*** (6.17)	20.25*** (3.75)	17.92*** (5.13)	-1.68 (10.94)	-1.45 (10.92)
$\ln(\text{GDP}(t-2)_G) - \ln(\text{GDP}(t-2)_H)$	-6.51* (3.45)	-7.82* (4.38)	-15.94*** (3.22)	-27.59*** (3.27)	0.91 (7.04)	1.42 (7.03)
Married	-1.36 (2.33)	-1.52 (2.92)	0.31 (0.75)	4.75*** (0.40)	-5.21* (2.68)	-5.28* (2.76)
Married living separated	-0.77 (2.72)	-0.81 (3.67)			-4.13 (2.95)	-4.16 (3.03)
Divorced	-2.57 (2.44)	-2.90 (3.01)	1.20 (2.66)	7.27*** (1.27)	-7.95*** (2.83)	-7.83*** (2.90)
Widowed	-2.12 (2.49)	-2.24 (3.00)			-6.19** (2.99)	-6.22** (3.04)
Employed	-0.88* (0.51)	-0.39 (0.56)	-1.53** (0.68)	-3.00*** (0.52)	-3.45*** (0.98)	-3.47*** (0.99)
Family at Home	6.00 (5.83)	4.72 (6.01)			11.02*** (0.95)	11.25*** (0.94)
Attended School in Germany	-25.35*** (1.09)	-25.01*** (1.19)				
Time in Germany	-0.63* (0.36)	-0.69* (0.41)	0.99** (0.46)	-0.33 (0.23)	-0.39 (0.73)	-0.38 (0.72)
Children?	-3.54 (2.40)	-2.94 (2.35)	-11.06*** (4.08)	-1.40*** (0.43)		
Aged 60 or older	3.48*** (0.73)	3.75*** (0.80)	3.26*** (0.71)	5.16*** (0.59)	3.71*** (1.24)	3.71*** (1.25)
Income	-0.00 (0.00)	-0.00 (0.00)	0.00 (0.00)	0.00*** (0.00)	0.00* (0.00)	0.00* (0.00)
Happiness	0.21** (0.09)	0.23** (0.11)	0.52*** (0.10)	0.37*** (0.08)	0.01 (0.13)	0.00 (0.13)
Disadvantage due to origin?	-0.75** (0.37)	-0.72* (0.43)			-1.22** (0.59)	-1.23** (0.59)
Writing German?	0.63 (0.86)	0.71 (0.92)			1.18 (1.86)	1.19 (1.86)
Speaking German?	-0.62 (2.60)	-0.44 (3.28)			-4.12* (2.48)	-4.12* (2.49)
Language Newspaper German?	0.93 (0.69)	0.96 (0.77)			1.84 (1.24)	1.83 (1.25)
Control Over Life		1.07* (0.58)				
Remitting			-2.40*** (0.61)			
Ever paid remittances?				-5.03*** (0.38)		
Medium low risktaker					-0.44 (0.59)	
Medium high risktaker					0.11 (0.72)	
High risktaker					-1.00 (1.14)	
Risktaker?						0.09 (0.56)
Constant	-11.14 (7.93)	-13.31 (10.72)	-35.27*** (10.19)	-41.35*** (2.27)	-3.59 (17.27)	-4.14 (16.97)
Bundesland FE	Yes	Yes	Yes	Yes	Yes	Yes
R ²	0.68	0.67	0.54	0.15	0.78	0.78
Number of Clusters	1950	1747	1749		1278	1278
Observations	12336	10572	17967	17976	4898	4898

Source: GSOEP, own calculations.

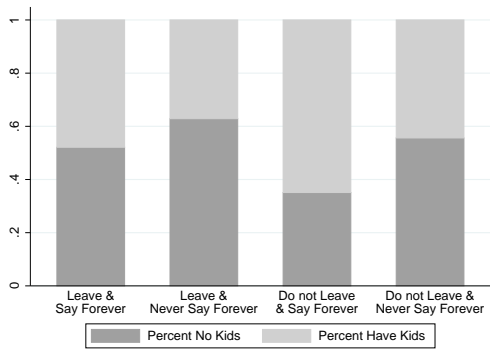
Note: *, **, *** indicates significance at the 10%, 5%, and 1% level, respectively. Standard errors in parentheses. The dependent variable is the difference between the intended return and the predicted realization. The columns that include the number of clusters, include individual fixed effects and those standard errors are clustered.



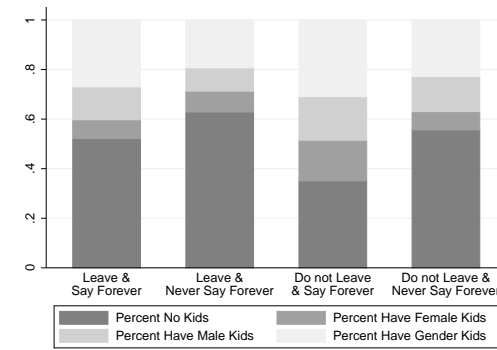
(a) Whole Sample



(b) Gender

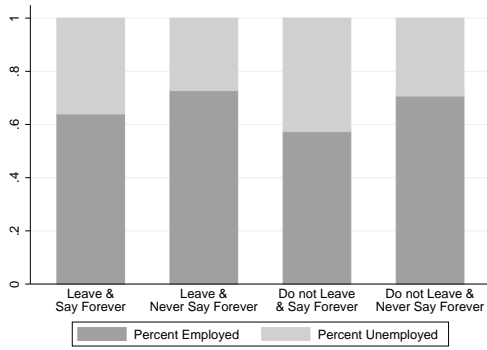


(c) Children

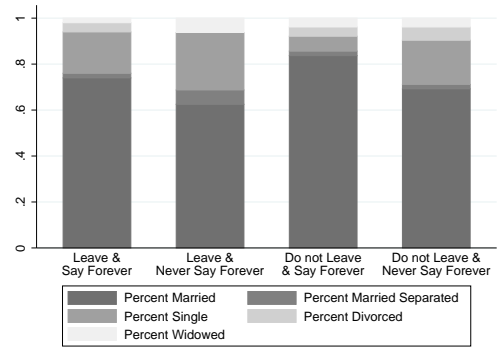


(d) Gender of the Children

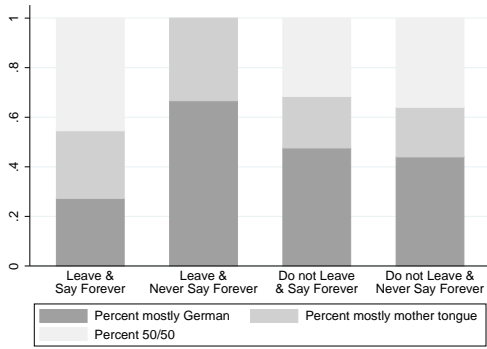
Figure 1: Descriptives



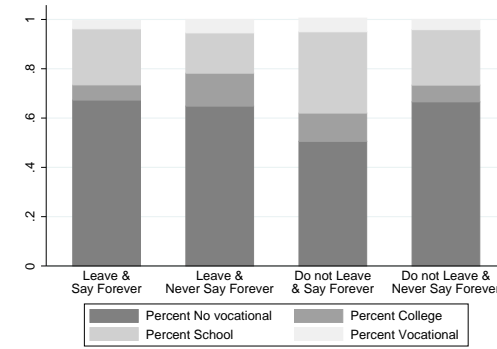
(a) Employed



(b) Marital Status

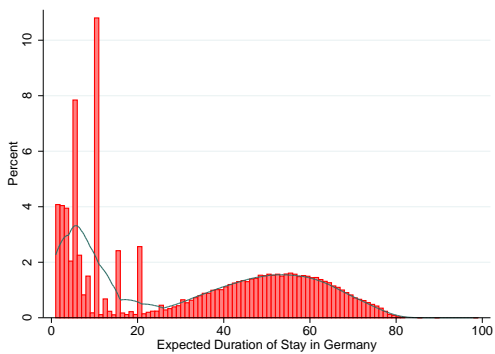


(c) Language Spoken at Home

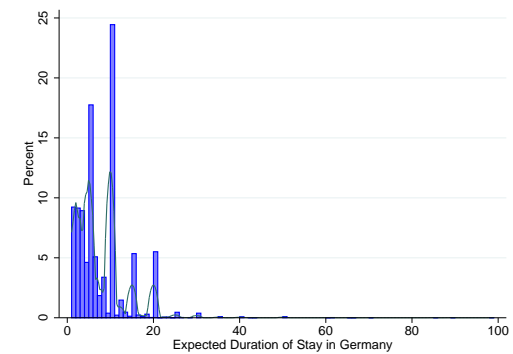


(d) Education

Figure 2: Descriptives

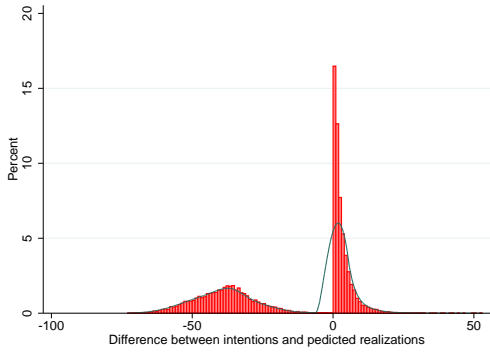


(a) Intended Duration of Stay for those who intend to stay forever = 100 - current age

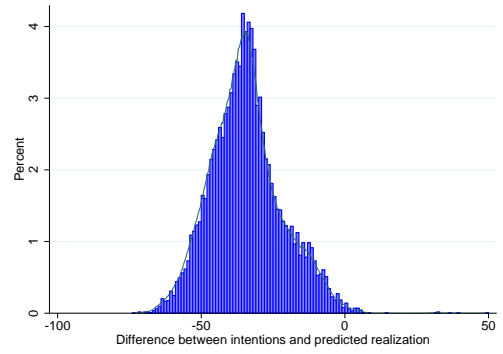


(b) Intended Duration of Stay without those who intend to stay forever

Figure 3: Expected Duration of Stay

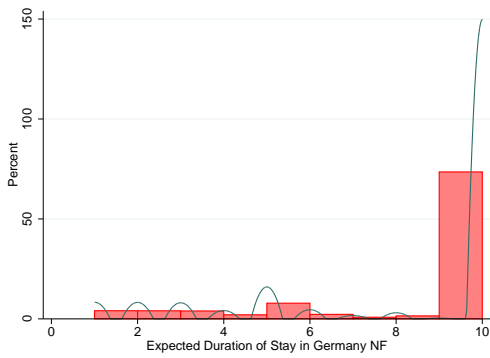


(a) Those who intend to stay forever maximal survival time = 100 - current age

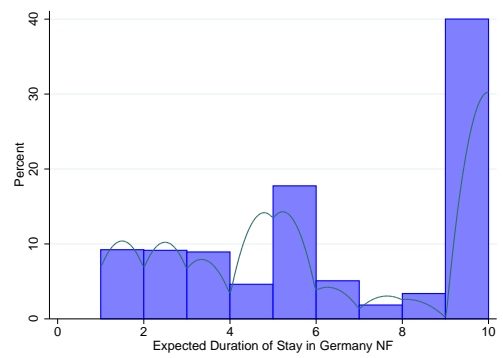


(b) Without those who intend to stay forever

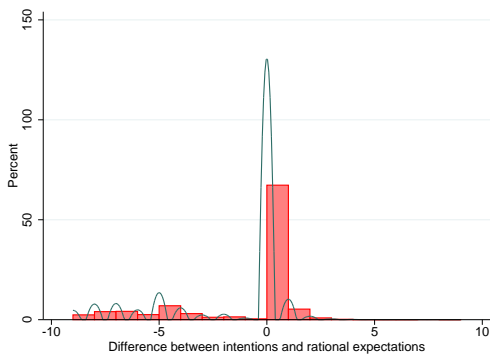
Figure 4: **Difference Between Intentions and predicted Realizations**



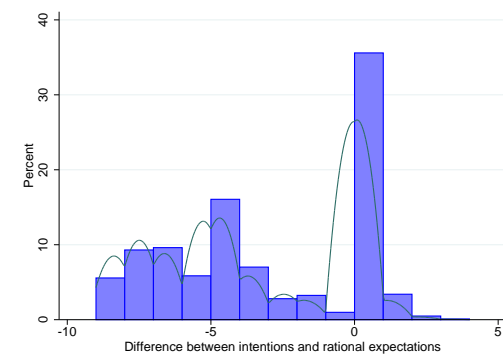
(a) Those who intended to stay forever, maximal survival time = 10



(b) Without those who intend to stay forever

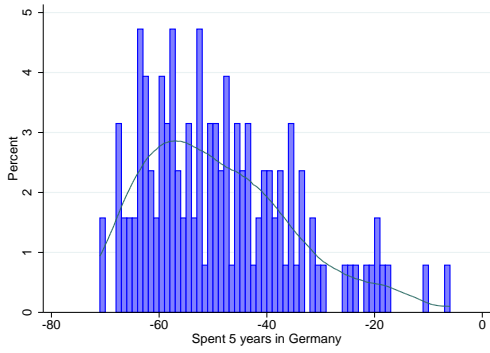


(c) Those who intended to stay forever, maximal survival time = 10



(d) Without those who intend to stay forever

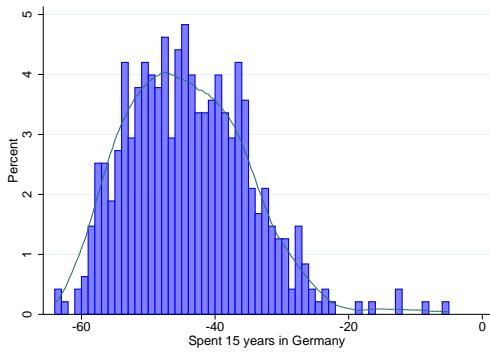
Figure 5: **“Narrow Framing”**



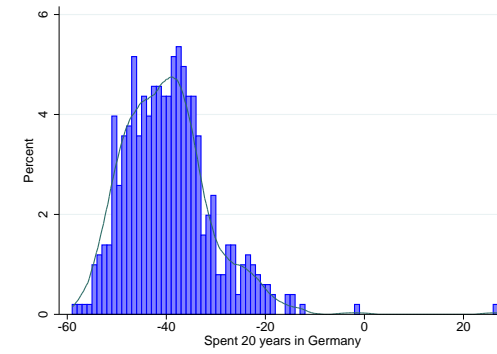
(a) Spent 5 years in Germany



(b) Spent 10 years in Germany

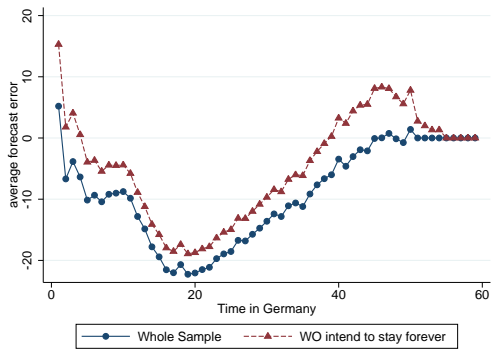


(c) Spent 15 years in Germany

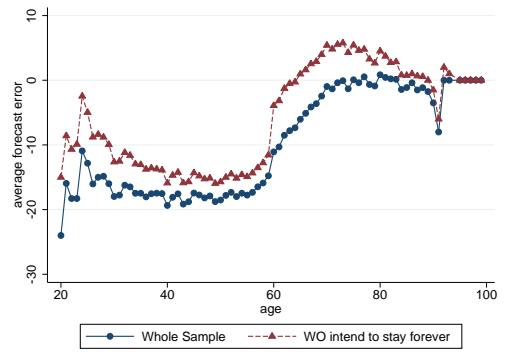


(d) Spent 20 years in Germany

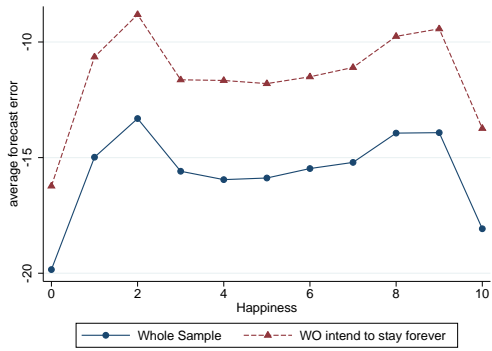
Figure 6: **Difference Between Intentions and predicted Return, Learning?**



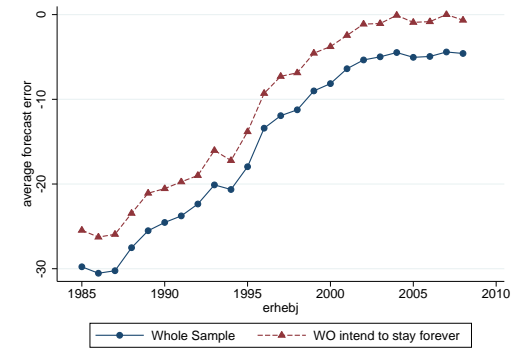
(a) Over time spent in Germany



(b) Over age at time of forecast



(c) Over Life Satisfaction at Time of Forecast



(d) Over time

Figure 7: Average Forecast Error

A Appendix

A.1 Possible Differences between Dustmann’s approach and the current approach

The “bioimmig.dta” file that is supplied by the SOEP and has already a panel form is used. In this data set the variable called “bistay” informs me about people’s intentions and the variable “bistayy” tells me how long they plan to stay in Germany. This structure seemed appealing, since - only the different information on the address log needed to be merged to the existing panel structure - here it was necessary to pay attention to the fact that the information about the return is at the household and not at the the individual level. Throughout the process of merging I came only across 3 people that split off their current household - their household moved out of Germany while they stayed in Germany.

One possible difference between my approach and Dustmann (2003a)’s approach could be that he constructed the panel structure himself, even though the information on the “bistay” variable should be the same whether I use the “bioimmig.dta” or whether I use the personal files for each year and append them. So far I haven’t had the time and opportunity to double check whether this will lead to the same results that I have or to the results that he has. It is definitely on my schedule to do a robustness check on my results with appending the different personal files.

Table A.1: **Return Frequency Dustmann (2003a) Table 1**

Year	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	Total
Return	163	67	59	74	53	41	29	35	36	43	35	31	22	688
Pct	4.74	1.95	1.72	2.15	1.54	1.19	0.84	1.02	1.05	1.25	1.02	0.90	0.64	20.02

Table A.2: **Intentions and Realization 1984 - 1997**

Return between 84 and 97			
Intended Return (84)	No	Yes	Total
No	705	59	764
Percentage	29.73	14.29	27.44
Yes	1666	354	2020
Percentage	70.27	85.71	72.56
Total	2371	413	2784

Reproducing Dustmann (2003a) with my sample

Source: GSOEP, own calculations.

Table A.3: **Intentions and Realizations 1984 - 1997 Dustmann (2003a) Table 2**

Return between 84 and 97			
Intended Return (84)	No	Yes	Total
No	665	98	763
Percentage	30.37	15.91	27.19
Yes	1525	518	2043
Percentage	69.63	84.09	72.81
Total	2190	616	2806

Dustmann (2003a) Table 2

A.2 Intentions and Residence Status

Different than in Table 2, the amounts off the diagonal are not as big anymore; out of those present in 1996 and saying that they want to return, only 38% never return. These numbers signal adjustment in contrast to Figures 1 and 2. The time horizon has become significantly smaller and between 60 to 70% of the respondents seem to predict their return correctly. Nevertheless the nearly 40% of the population that do not predict their future correctly do not present simple noise. Again the possible source of confound that individuals do not tell the truth about their residence status is not confirmed. Out of those that have a limited residence status, 65% claim that they want to stay in Germany forever.

Table A.4: **Intentions and Realization 1996 - 2009**

Intended Return (96)	Return between 96 and 09		Total
	No	Yes	
No	870	35	905
Percentage	60.80	30.97	58.61
Yes	561	78	639
Percentage	39.20	69.03	41.39
Total	1431	113	1544

Source: GSOEP, own calculations.

Table A.5: **Desire to Return versus Residence Status 1996**

Desire to Return	Residence Status		
	Unlimited	Limited	Total
Within 12 Months	0	2	2
(Percentage)	0.00	2.38	1.40
After One Year	24	27	51
Percentage	40.68	32.14	35.66
Stay in Germany Forever	35	55	90
Percentage	59.32	65.48	62.94
Total	59	84	143

Source: GSOEP, own calculations.

Table A.6: Socioeconomic Differences for Stayers

Variable	Right Censored Obs			Attritors			t-stat
	Mean	SD	N	Mean	SD	N	
Male	0.54	0.50	1209	0.48	0.50	2682	(-3.52)***
Age at Migration	30.42	11.14	1209	29.87	10.43	2682	(-1.49)
$\ln(\text{GDP}_G) - \ln(\text{GDP}_H)$	1.68	1.25	1192	1.69	1.02	2646	(0.29)
Married	0.80	0.40	1209	0.57	0.50	2355	(-14.38)***
Married living separated	0.02	0.15	1209	0.02	0.14	2355	(-0.53)
Divorced	0.07	0.26	1209	0.04	0.20	2355	(-4.10)***
Widowed	0.06	0.25	1209	0.05	0.21	2355	(-2.25)
Employed	0.50	0.50	1209	0.54	0.50	2681	(2.39)*
Family at Home	0.31	0.46	1203	0.14	0.35	2673	(-12.05)***
Spouse at Home	0.01	0.09	1209	0.02	0.15	2682	(3.15)***
Attended School in Germany	0.04	0.20	1189	0.03	0.16	2643	(-2.56)*

Source: GSOEP, own calculations.

Note: The t-statistics test for the significance of the difference between right censored individuals and those that disappear before 2008. For each individual the last point in time where information is provided in the data set is taken to get the different means.

Table A.7: Country of Origin

Country	Frequency	Percent
Turkey	9,670.0	22.9
Ex-Yugoslavia	4,357.0	10.3
Greece	3,824.0	9.1
Italy	4,996.0	11.9
Spain	2,326.0	5.5
Austria	518.0	1.2
France	304.0	0.7
Benelux	75.0	0.2
Great Britain	251.0	0.6
USA	186.0	0.4
Switzerland	127.0	0.3
Romania	1,170.0	2.8
Poland	2,650.0	6.3
Iran	139.0	0.3
Hungary	224.0	0.5
Portugal	80.0	0.2
Bulgaria	114.0	0.3
Czech Republic	287.0	0.7
Russia	2,012.0	4.8
Philippines	156.0	0.4
Kazakhstan	1,561.0	3.7
Albania	80.0	0.2
Kirgistan	88.0	0.2
Ukraine	415.0	1.0
Tadzhikistan	67.0	0.2
Vietnam	67.0	0.2
Netherlands	236.0	0.6
Croatia	1,439.0	3.4
Bosnia Herzegovina	848.0	2.0
Macedonia	164.0	0.4
Slovenia	192.0	0.5
Kosovo Albania	163.0	0.4
Eastern Europe	1,578.0	3.7

Table A.8: Difference between the Intentions and the 'rational' Expectations 1984 - 1996

	85	86	87	88	89	90	91	92	93	94	95	96
Male	-1.16 (1.48)	-3.82** (1.48)	-3.14** (1.48)	-3.27** (1.50)	-1.43 (1.59)	-0.19 (1.61)	-2.62 (1.60)	-0.79 (1.63)	0.49 (1.65)	-2.42 (1.65)	-2.78* (1.67)	-1.01 (1.20)
Age at Migration	0.31*** (0.09)	0.40*** (0.09)	0.39*** (0.09)	0.19* (0.10)	0.27*** (0.11)	0.08 (0.11)	-0.05 (0.12)	0.06 (0.13)	0.21 (0.13)	0.17 (0.14)	0.06 (0.15)	0.33*** (0.08)
ln(GDP _G)-ln(GDP _H)	-4.04*** (0.94)	-4.04*** (0.92)	-3.80*** (0.90)	-2.16** (0.93)	-2.34** (0.96)	-2.97*** (0.97)	-1.01 (0.96)	-3.98*** (0.99)	-2.41** (1.01)	-0.63 (1.01)	-3.59*** (1.03)	2.40*** (0.71)
Employed	-9.21*** (2.11)	-7.19*** (2.25)	-8.47*** (2.26)	-6.38*** (2.15)	-5.22** (2.53)	-8.85*** (2.61)	-3.96 (2.60)	-3.20 (2.65)	-4.60* (2.56)	-1.83 (2.77)	1.51 (2.80)	-4.55** (2.04)
Spouse at Home	2.77 (3.08)	6.12** (3.03)	6.26** (3.34)	3.54 (3.34)	6.93** (3.14)	6.50* (3.59)	6.50* (3.73)	2.25 (4.10)	-2.43 (4.02)	2.16 (4.44)	2.16 (4.63)	-5.05 (5.15)
Attended School in Germany	2.26 (3.57)	3.18 (3.40)	3.28 (3.32)	7.35** (3.71)	2.90 (3.58)	-2.63 (3.78)	-3.38 (3.67)	2.30 (3.81)	4.09 (3.75)	5.08 (4.00)	9.22** (4.14)	1.92 (3.78)
Time in Germany	-4.54*** (1.34)	-3.84*** (1.30)	-4.70*** (1.47)	-5.26*** (1.71)	-1.06 (1.95)	-1.59 (2.24)	-4.43* (2.40)	-3.99 (2.70)	-2.85 (2.92)	-6.48** (3.00)	-1.84 (3.74)	-0.72 (0.80)
Time in Germany ²	0.33*** (0.09)	0.26*** (0.09)	0.29*** (0.09)	0.29*** (0.10)	0.07 (0.11)	0.10 (0.12)	0.24** (0.12)	0.22* (0.13)	0.15 (0.13)	0.29** (0.13)	0.10 (0.16)	-0.01 (0.04)
Time in Germany ³	-0.01*** (0.00)	-0.00** (0.00)	-0.00** (0.00)	-0.00** (0.00)	-0.00 (0.00)	-0.00 (0.00)	-0.00 (0.00)	-0.00 (0.00)	-0.00 (0.00)	-0.00 (0.00)	-0.00 (0.00)	0.00 (0.00)
Income	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
Happiness	0.84*** (0.29)	0.89*** (0.29)	0.60*** (0.30)	0.66** (0.33)	0.91*** (0.34)	0.20 (0.35)	0.69* (0.36)	0.37 (0.37)	0.42 (0.36)	-0.37 (0.38)	0.17 (0.39)	-0.11 (0.32)
Aged 60 or older	3.15 (3.28)	3.09 (3.15)	3.48 (2.91)	6.29** (2.92)	5.78** (2.84)	9.82*** (2.79)	10.83*** (2.68)	10.33*** (2.76)	7.30*** (2.69)	7.34*** (2.66)	9.25*** (2.65)	0.37 (2.15)
Married	11.26 (13.52)	11.26 (13.52)	18.64** (8.93)	10.87 (10.17)	8.87 (10.21)	16.32* (8.62)	26.17*** (20.54)	20.64*** (20.95)	14.67** (20.95)	18.19*** (20.41)	19.02*** (19.93)	-3.48 (3.94)
Divorced												
Widowed												
Married living separated												
Family at Home												
Constant	-22.71** (9.52)	-43.19*** (8.71)	-31.93*** (9.61)	-10.24 (13.44)	-46.20*** (12.90)	-28.44* (15.06)	-13.04 (16.81)	-5.59 (19.51)	-0.84 (24.74)	33.65 (25.85)	-8.23 (31.21)	-2.37 (20.59)
Country Region	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Bundesland FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R ²	0.15	0.18	0.15	0.16	0.17	0.11	0.15	0.13	0.15	0.15	0.16	0.23
F-stat.	9.13	9.77	7.39	7.12	7.85	5.06	6.14	4.79	5.16	5.17	4.95	8.37
p-val.	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
N	1367	1254	1195	1115	1094	1050	988	938	896	827	755	1252

Source: GSOEP, own calculations.

Note: *** indicates significance at the 10%, 5%, and 1% level, respectively. Standard errors in parentheses. The dependent variable is the difference between the intended return and the 'rational' expectation.

Table A.9: Difference between the Intentions and the 'rational' Expectations 1997 - 2008

	97	98	99	00	01	02	03	04	05	06	07	08
Male	0.80 (1.44)	0.84 (1.37)	-0.37 (1.27)	0.23 (1.26)	-1.02 (1.04)	-0.67 (0.98)	-0.32 (0.97)	-0.72 (0.98)	0.04 (0.99)	0.29 (1.04)	0.09 (1.04)	-0.46 (1.08)
Age at Migration	0.23** (0.09)	0.16** (0.09)	0.16** (0.08)	0.25*** (0.08)	0.18*** (0.07)	0.19*** (0.07)	0.07 (0.07)	0.03 (0.07)	0.05 (0.07)	0.06 (0.07)	-0.01 (0.07)	0.12 (0.08)
ln(GDP _C)-ln(GDP _H)	1.75** (0.75)	1.91*** (0.72)	1.88*** (0.66)	2.13*** (0.65)	2.38*** (0.52)	2.05*** (0.49)	2.41*** (0.49)	1.80*** (0.49)	1.09** (0.51)	1.91*** (0.53)	2.09*** (0.54)	2.85*** (0.57)
Married	-11.31*** (3.23)	-8.87*** (3.23)	(0.66) (3.07)	(0.65) (3.10)	(0.52) (2.22)	(0.49) (2.22)	(0.49) (2.44)	(0.49) (2.44)	(0.51) (2.68)	(0.53) (2.81)	(0.54) (2.97)	(0.57) (3.19)
Married living separated	-14.20** (5.96)	-4.83 (6.04)	-3.33 (5.96)	0.14 (5.77)	-5.65 (4.30)	-0.27 (4.30)	8.21** (3.91)	-2.30 (4.02)	-2.16 (3.82)	-12.85*** (4.04)	-7.85* (4.11)	-6.48 (4.45)
Divorced	-5.34 (4.55)	-8.11* (4.39)	-3.75 (4.00)	-6.12 (3.95)	1.60 (2.94)	0.34 (2.93)	6.99** (3.05)	-2.21 (3.00)	1.59 (3.25)	-4.68 (3.32)	-3.82 (3.42)	-9.24** (3.63)
Widowed	-14.08*** (5.08)	-9.79* (5.00)	-3.29 (4.37)	-4.19 (4.60)	1.07 (3.51)	-2.44 (3.36)	3.49 (3.47)	-4.66 (3.44)	0.99 (3.45)	-6.83* (3.67)	-6.76* (3.70)	-9.02* (3.82)
Employed	1.58 (2.22)	-0.81 (1.96)	11.53*** (3.20)	-2.14 (1.89)	-0.09 (1.37)	-2.31* (1.28)	-4.96*** (1.32)	-3.46*** (1.42)	-3.74** (1.42)	-6.24*** (1.44)	-3.74** (1.46)	-3.95** (1.54)
Family at Home	2.86 (2.34)	4.13* (2.11)	4.06** (1.82)	3.11* (1.80)	2.06 (1.26)	1.11 (1.18)	2.01* (1.14)	1.04 (1.15)	2.21* (1.16)	3.23*** (1.22)	0.43 (1.20)	1.48 (1.24)
Spouse at Home	-0.42 (5.59)	0.13 (5.97)	-6.33 (5.19)	0.28 (5.40)	-0.61 (5.33)	-3.19 (5.01)	-4.39 (5.54)	-2.77 (5.21)	-2.05 (5.40)	2.09 (5.41)	6.63 (5.64)	3.17 (5.51)
Attended School in Germany	-4.14 (3.79)	2.62 (3.68)	5.56* (3.33)	-3.22 (3.25)	-0.92 (2.70)	2.45 (2.70)	2.27 (2.48)	0.23 (2.51)	0.97 (2.66)	1.27 (2.59)	1.27 (2.57)	3.07 (2.59)
Time in Germany	-2.07** (0.89)	-0.33 (0.91)	-0.34 (0.82)	1.07 (0.86)	0.89* (0.53)	-0.29 (0.50)	0.72 (0.52)	-0.22 (0.51)	-0.58 (0.53)	0.02 (0.54)	0.59 (0.63)	0.33 (0.68)
Time in Germany ²	0.07 (0.05)	-0.01 (0.04)	-0.01 (0.04)	-0.07 (0.04)	-0.05*** (0.03)	0.01 (0.02)	-0.04 (0.02)	0.00 (0.02)	0.01 (0.02)	-0.01 (0.02)	-0.03 (0.03)	-0.02 (0.03)
Time in Germany ³	-0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00** (0.00)	0.00 (0.00)	0.00* (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
Income	0.00 (0.00)	0.00 (0.00)	-0.00 (0.00)	0.00 (0.00)	-0.00 (0.00)	-0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
Happiness	0.02 (3.77)	-0.20 (2.87)	-0.15 (3.32)	0.12 (3.33)	-0.05 (2.92)	-0.26 (2.28)	0.11 (2.26)	-0.06 (2.26)	0.58** (2.25)	0.50* (2.27)	0.32 (2.28)	0.36 (2.28)
Aged 60 or older	3.72 (2.38)	2.87 (2.31)	1.87 (2.06)	1.04 (2.06)	1.54 (1.80)	0.92 (1.77)	2.29 (1.75)	4.75*** (1.79)	3.34* (1.78)	2.86 (1.87)	4.45** (1.89)	2.44 (1.94)
Disadvantage due to origin?	-4.18*** (1.29)	-1.43 (1.25)	-4.44*** (1.15)	-2.33** (1.18)	0.34 (1.00)	-1.02 (0.98)	-1.84* (0.96)	-1.87* (0.98)	-0.31 (1.01)	1.16 (1.02)	-1.00 (1.02)	0.05 (1.10)
Writing German?	-0.31 (1.85)	-3.12* (1.82)	1.58 (1.68)	1.26 (1.68)	-1.27 (1.54)	-1.59 (1.52)	13.32* (1.52)	13.32* (1.52)	0.63 (1.72)	-2.13 (1.78)	-4.52*** (1.81)	-5.75*** (1.94)
Speaking German?	5.89 (6.57)	1.01 (5.41)	7.36 (5.80)	-1.30 (5.31)	0.11 (4.45)	4.25 (3.79)	4.69 (3.60)	3.72 (3.51)	0.87 (4.13)	-2.35 (4.62)	-2.50 (5.23)	-0.60 (5.12)
Language Newspaper German?	6.42*** (1.60)	8.78*** (1.55)	6.48*** (1.50)	5.27*** (1.56)	5.35*** (1.32)	9.10*** (1.32)	3.77*** (1.30)	4.29*** (1.33)	3.64*** (1.42)	4.60*** (1.42)	2.47* (1.45)	7.62*** (1.56)
Constant	-19.92* (11.57)	0.60 (14.99)	-29.83* (15.30)	-24.87* (15.04)	-17.52* (9.32)	-15.58* (8.31)	-27.94*** (10.60)	-21.98* (11.99)	-15.92 (10.10)	-16.60 (13.10)	-7.30 (17.01)	-12.28 (16.62)
Country Region	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Bundesland FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R ²	0.28	0.25	0.27	0.22	0.19	0.20	0.17	0.16	0.17	0.19	0.16	0.20
F-stat.	7.38	6.38	6.80	5.31	5.72	5.88	4.61	4.23	4.22	4.49	3.36	3.88
p-val.	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
N	920	960	937	938	1265	1194	1186	1171	1075	979	907	805

Source: GSOEP, own calculations.

Note: *, **, *** indicates significance at the 10%, 5%, and 1% level, respectively. Standard errors in parentheses. The dependent variable is the difference between the intended return and the 'rational' expectation.