

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

By Gerard J. van den Berg and Michèle A. Weynandt*

Abstract

This paper explores the difference between intentions and realizations in return migration with the help of a duration model. Using the SOEP the results lend 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. Along these lines we find that migrated individuals systematically underestimate the length of their stay in the receiving country. 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, D84, F22

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 (see e.g; Read/van Leeuwen, 1998; Gilbert/Michael/Gill/Wilson, 2002; Nisbett/Kanouse 1968). The reader is presented with evidence of a forecasting error and convincing statistics proving that it is not just simple noise. Loewenstein/O’Donoghue/Rabin (2003) have defined the concept of projection bias as the suggestion that people understand

* We thank the editor Nicolas Ziebarth and an anonymous Referee for very helpful comments and suggestions. 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 and SOEP 2012, for many helpful comments on an earlier version of this paper.

the qualitative nature of changes in their tastes, but underestimate the magnitude of these changes.¹

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/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 on.²

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.³ 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 (SOEP) is used for the analysis as individuals provide information on their return intentions.⁴ Using a duration model an expression for the predicted return realization (expected duration of the stay in Germany for the current analysis) is inferred. This predicted return will be compared to the respondents intentions and will be regressed on different sets

¹ Conlin/O'Donoghue/Vogelsang (2007); Levy (2009); Acland/Levy (2010) are papers that are able to pin down the exact value of the projection bias parameter.

² Individuals who 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.

³ See Sjaastad (1962) for a first formulation of the decision to move. For a more thorough review on the return migration decision, see Borjas/Bratsberg (1996) who generalize Borjas (1991). Other relevant studies are; Dustmann (2003b); Dustmann (2003a); Dustmann/Weiss (2007).

⁴ To get a thorough overview of the data, we refer you to Wagner/Frick/Schupp (2007). The Question asked is: Do you want to stay in Germany forever? This can be answered by yes or no. If the respondents answer no, they are asked how long do you want to remain in Germany, which can be answered by a year at most and a few more years. If the individual then says a few more years, they are asked how many years? And they are supposed to give an estimated amount of years. E.g. if the individual responds 5 it is supposed to mean she wants to return in 5 years.

of socio-economic variables. These regression results highlight the driving factors of the difference between return intentions and return realizations.⁵

A first important finding is that people's intentions exert bunching which already points towards a simplifying heuristic. 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 error is therefore mostly negative but decreases the longer one stayed in Germany and the older one gets.⁶ Using pooled OLS we are also able to highlight 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.

The clear understanding of the difference between intentions 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 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 presents the duration model. Section 3 presents the results, while section 4 concludes.

2. 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 at t or the return probability, where 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 would 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

⁵ Please be aware that we are not claiming a causality of the results. We are only interested in the driving factors of the forecast error.

⁶ The 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.

an expression for the return realization which will be inferred through duration analysis.

GDP is a good indicator to compare countries and as mentioned in Borjas/Bratsberg (1996) or Dustmann/Weiss (2007) 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.⁷

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 (Sjaastad, 1962). More formally, if one decides to move in 2005,

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

needs to hold.⁸ This can also be rewritten in terms of the hazard rate in 2005, such that:

$$(2) \quad P\left(\epsilon < \sum_{t=2005}^K \frac{1}{(1+r)^t} (E[U(X_H(t))] - E[U(X_G(t))]) - c\right) \\ = \Phi\left(\frac{\sum_{t=2005}^K \frac{1}{(1+r)^t} (E[U(X_H(t))] - E[U(X_G(t))]) - c}{\sigma_\epsilon}\right).$$

This expression can also be rewritten to get an expression for the hazard rate each year. Since we are ultimately interested in the expected duration of a stay, the duration framework allows us to write:

$$(3) \quad y(0) = E(T|x_0, \text{expectations of future path of } x(t)) \\ = \int_0^\infty \left[\exp\left(-\int_0^\infty \theta(u|x(u), x_0) du\right) \right] dz$$

⁷ The GDP levels for the different countries are from Angus Maddison (<http://www.ggdc.net/MADDISON/oriindex.htm>) but are only available until 2008, which forces the drop of the year 2009 and leaves us with the period 1984–2008.

⁸ $X(t)$, are covariates that we control for. c represents the cost of moving, K is the maximum survival time for which we assume that it equals the expected lifetime duration approximated by $100 - \text{current age}$ in the empirical part. r is the interest rate and ϵ is an error term. $U()$ represents the utility function. The subscript G stands for Germany, and the subscript H stands for Home.

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"?

In the empirical part, we are going to estimate a discrete time complementary log log model with a third order polynomial in time, which allows to infer the predicted return.⁹

3. Results

The first impression of a bias becomes visible when we compare the actual and the intended return. Looking at Table 1 about 70 percent of non-returners had the intention to return to their home country in 1984. When evaluating Table 1 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 SOEP surveys people annually thereby not allowing to account for people who migrate and return within a year.¹⁰ For the rest of the paper we consider only the migrants who are already in Germany and present in the SOEP. Furthermore we only consider adults who are older than 18 years in order to only include those individuals who take the return decision themselves. As the use of the GDP Data forced the drop of the year 2009, we are left with 25 years for the analysis (1984–2008).¹¹

Before analyzing the differences between the intentions and the predicted return, let us look at the individuals intentions and what may be driving factors

⁹ The results for the cloglog model are not reported here, but are available from the authors upon request. We control for a third order polynomial in time, gender, age at migration, difference in GDP between Germany and the home country, marital status, whether or not one has family at home, a spouse at home, attended school in Germany, GDP growth in Germany and at home, income, whether or not one is above 60 and country of origin. The dependent variable is 1 if the individual has left Germany and 0 if we do not observe the respondent leave. Whether or not an individual has left is inferred from the address log of the SOEP.

¹⁰ These individuals do not play an important role for the analysis of the underestimation of the trip duration.

¹¹ Table 1 only includes statistics for those individuals who answered the intention to return question in 1984. In table 2 the number of individuals is higher and it depends upon the specification exactly how many. For the cloglog models we had 3152 individuals where 574 durations until re-migration are not right censored.

Table 1
Intentions and Realization 1984–2009

Intended Return (84)	Return between 84 and 09		
	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.

of changes in these intentions. Figure 1 plots the Intended Duration of Stay, in figure 1a we imputed the intended duration for those who wanted to stay forever as 100 – their current age, while in figure 1b 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 at work when individuals form their intentions.¹² Figure 2 plots the difference between the intentions and the predicted return. Again figure 2a plots the difference for the whole sample where for those who intended to stay forever we imputed their maximal survival time as 100 – their current age. Figure 2b plots the differ-

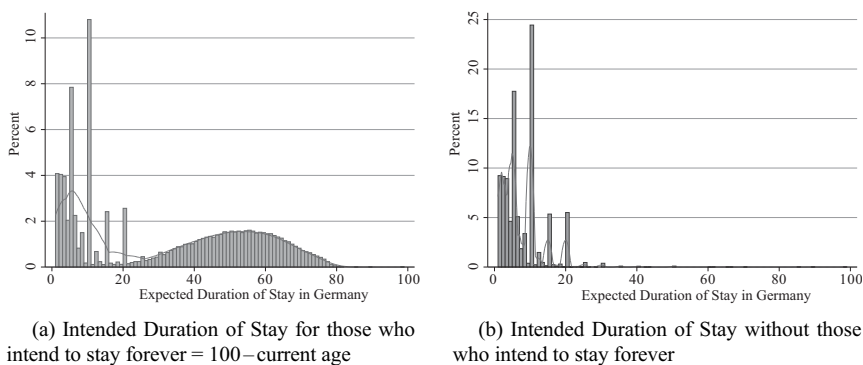


Figure 1: Intended Duration of Stay

¹² We also ran regressions on what drives the changes in intentions, these results are available from the authors upon request.

ence for the reduced sample where we leave those out who intend to stay forever. Figure 2a gives us hope that there seem to be many people predicting the duration of their stay correctly, but when we take those out who intend to stay forever (which are those that we imposed on how long they remain in Germany – figure 2b) practically no remaining individual has a correct prediction. Figure 2b 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.

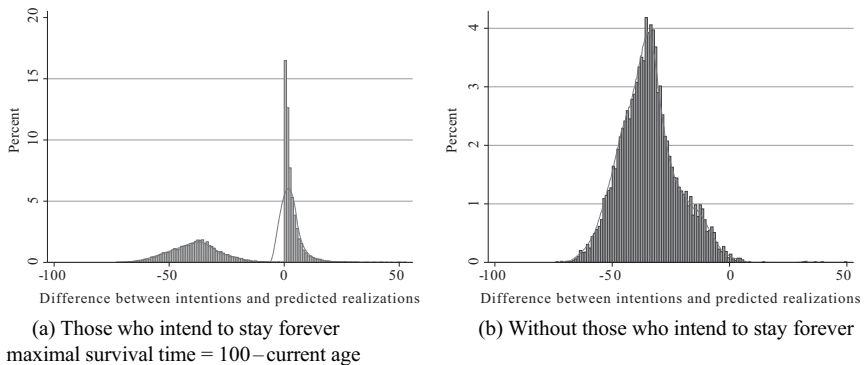


Figure 2: Difference Between Intentions and Predicted Realizations

Figure 3 and 4 then plot the average forecast error (equivalent to the average difference) over different time specifications. Figure 3a plots the average forecast error over time spent in Germany, and what was not obvious before now seems to become relevant. 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 estimating your actual chance of leaving and so you seem to be more accurate with your forecast.

Figure 3b 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 not 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. Figure 4 then plots the average forecast error over time. We included this result in order to show that there do not seem to be relevant macro shocks that may drive our results.

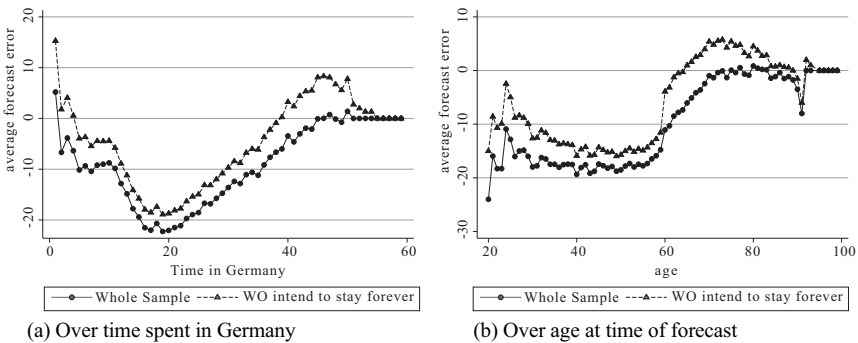


Figure 3: Average Forecast Error

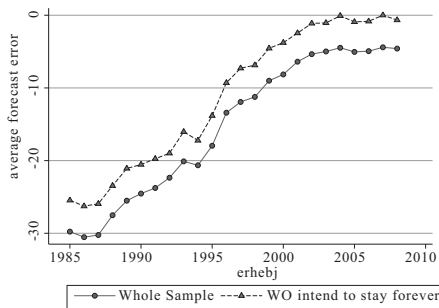


Figure 4: Average Forecast Error

Table 2 then finally takes a closer look at the differences between the intentions and the predicted return. In table 2 columns (3), (5) and (6) include individual fixed effects where the standard errors are also clustered at the individual level. The OLS results in columns (1), (2) and (4) are shown just for comparison.¹³ The coefficient on ‘attended school in Germany’ changes sign and

¹³ The 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

magnitude as soon as we include individual fixed effects and clustering at the individual level. The coefficient is significant in the fixed effects regressions at the 1 % level and negative while in the OLS results it is positive and not always significant. 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 who feel a disadvantage underestimate the duration of their stay by more than those who do not feel disadvantaged. Having attended school in Germany is one of the surprising coefficients since it increases the difference when we focus on the specifications which include individual fixed effects in table 2 columns (3), (5) and (6). The coefficient on the happiness variable (general life satisfaction) is positive and significant at the 5 % level, meaning that the higher the general life satisfaction the smaller the difference. This points into the right direction. Those individuals who seem to be happier, are also those who make less mistakes when trying to predict their future.

may be time invariant, yearly OLS results are available upon request from the authors. The number of observations between columns (1) and the rest differs drastically due to non-response. Some individuals never reply to the questions referring to the German language. Just as much as some individuals never reply to the behavioral variables.

Table 2

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)		
$\ln(\text{GDP}_G) - \ln(\text{GDP}_H)$	-1.31 (2.62)	-17.97*** (3.69)	-4.75 (3.63)	-6.18 (23.26)	-4.90 (3.69)	-5.04 (3.68)
$\ln(\text{GDP}(t-1)_G) - \ln(\text{GDP}(t-1)_H)$	19.36*** (4.29)	18.34*** (5.67)	13.23*** (4.30)	13.55 (23.68)	13.84*** (4.54)	13.94*** (4.54)
$\ln(\text{GDP}(t-2)_G) - \ln(\text{GDP}(t-2)_H)$	-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: SOEP, 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 (3), (5) and (6), include individual fixed effects and there the standard errors are clustered at the individual level.

4. Conclusion

This 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. Therefore their prediction gets better and more accurate. This goes along with findings of Smith/Taylor/Sloan (2001) who found that longevity expectations are consistently linked to subsequent observed mortality. Another interpretation of the results leads towards Kahneman (2011) description of 'what you see is all there is' (WYSIATI). As shortly mentioned, 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 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.

Furthermore, we aim to explore whether differences between intentions and actual return migration are systematically related to behavioral traits of the respondents. This includes indicators of self-control and well-being. Obviously such investigations can only be carried out with rich longitudinal survey data such as the SOEP.

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./Levy, M.* (2010): Habit Formation and Naivete, and Projection Bias in Gym Attendance, mimeo.
- Borjas, G. J.* (1991): Immigration and Self-Selection in: J. M. Abowd/R. B. Freeman (eds.), *Immigration, trade and the labor market*, Chicago, IL.
- Borjas, G. J./Bratsberg, B.* (1996): Who leaves? The outmigration of the foreign-born, *Review of Economics and Statistics* 78(1), 165–176.
- Card, D./Dustmann, C./Preston, I.* (2012): Immigration, wages and compositional amenities, *Journal of the European Economic Association* 10 (1), 78–119.
- Conlin, M./O'Donoghue, T./Vogelsang, T. J.* (2007): Projection Bias in Predicting Catalog Orders, *American Economic Review* 97(4), 1217–1249.
- 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./Weiss, Y.* (2007): Return Migration: Theory and Empirical Evidence from the UK, *British Journal of Industrial Relations* 45(2), 236–256.
- Gilbert, D./Michael, T./Gill, J./Wilson, T. D.* (2002): The Future is Now: Temporal Correction in Affective Forecasting, *Organizational Behavior and Human Decision Processes*, LXXXVIII, 430–444.
- Kahneman, D.* (2011): *Thinking, Fast and Slow*.
- Levy, M.* (2009): *An Empirical Analysis of Biases in Cigarette Addiction*, mimeo.
- Loewenstein, G./O'Donoghue, T./Rabin, M.* (2003): Projection Bias in Predicting Future Utility, *Quarterly Journal of Economics* 118(4), 1209–1248.
- Nisbett, R. E./Kanouse, D. E.* (1968): Obesity, Hunger, and Supermarket Shopping Behavior, *Proceedings of the Annual Convention of the American Psychological Association* III, 683–684.
- Read, D./van Leeuwen, B.* (1998): Predicting Hunger: The Effects of Appetite and Delay on Choice, *Organa* LXXXVI, 189–205.
- Sjaastad, L. A.* (1962): The Costs and Returns of Human Migration, *Journal of Political Economy*, 70(5), 80–93.
- Smith, K. V./Taylor, D. H./Sloan, F. A.* (2001): Longevity Expectations and Death: Can People Predict Their Own Demise?, *The American Economic Review* 91(4), 1126–1134.

Wagner, G./Frick, J./Schupp, J. (2007): The German Socio-Economic Panel Study (SOEP) – Scope, Evolution and Enhancements, Schmollers Jahrbuch 127(1), 139–169.