IZA Institute of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 14155

Why Do Migrants Stay Unexpectedly? Misperceptions and Implications for Integration

Marc Kaufmann
Joël Machado
Bertrand Verheyden

MARCH 2021

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 14155

Why Do Migrants Stay Unexpectedly? Misperceptions and Implications for Integration

Marc Kaufmann
CEU

Joël Machado
LISER and IZA

Bertrand Verheyden
LISER

MARCH 2021

Any opinions expressed in this paper are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but IZA takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The IZA Institute of Labor Economics is an independent economic research institute that conducts research in labor economics and offers evidence-based policy advice on labor market issues. Supported by the Deutsche Post Foundation, IZA runs the world's largest network of economists, whose research aims to provide answers to the global labor market challenges of our time. Our key objective is to build bridges between academic research, policymakers and society.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Why Do Migrants Stay Unexpectedly? Misperceptions and Implications for Integration*

Empirical evidence suggests that a large proportion of immigrants who initially intended to stay temporarily in the destination country end up staying permanently, which may lead to suboptimal integration. We study systematic causes of unexpected staying that originate in migrant misperceptions. Our framework contains uncertainty about long-term wages, endogenous integration and savings in the short term, and return migration in the long term. We identify necessary and sufficient conditions on misperceptions that lead migrants to overestimate their probability of return migration, independently of their characteristics. We show that these conditions involve pessimism about the destination country, either in terms of short-term utility, of long-term utility, or of wage prospects. We then highlight specific behavioural biases that give rise to such forms of pessimism. Using the German Socio-Economic Panel, we find that relatively higher pessimism at arrival about future utility and wages is associated with migrants staying unexpectedly ex post.

JEL Classification: F22, D91, J61

Keywords: migrant integration, return intentions, unexpected staying, misperceptions, pessimism, GSOEP

Corresponding author:
Bertrand Verheyden
Maison des Sciences Humaines 11
Porte des Sciences
L-4366 Esch-sur-Alzette / Belval
Luxembourg

E-mail: bertrand.verheyden@liser.lu

* The authors thank Michel Beine, Michal Burzynski, Frédéric Docquier, Martín Fernández-Sánchez, Christina Gathmann, Thomas Gautier, Thomas Liebig, Tobias Müller, Hend Shehata, Felix Stips, Michel Tenikue, Gaspare Tortorici and participants at LISER's internal seminar and the 10th annual conference on Immigration in OECD countries for useful comments and suggestions. Any remaining errors are the sole responsibility of the authors. The authors gratefully acknowledge financial support by the Luxembourg National Research Fund (FNR), Project ID: 11343831.

1 Introduction

Among the innumerable choices we make, a few stand out as forks in the road: whom to have a relationship with, whether to have kids, which job to take – as well as the decision where to live. Large numbers of people migrate each year in the hopes of improving their lot. Despite the high stakes of migrating, substantial survey evidence suggests that migrants systematically mispredict their length of stay (Schoorl (2011) for the Netherlands, Adda et al. (2006); van Baalen and Muller (2008); van den Berg and Weynandt (2013) for Germany, Agyeman (2011) for Italy and Spain, Alberts and Hazen (2005) for the US, and Achenbach (2017) for Japan). Specifically, significant proportions of migrants end up staying longer in the destination country than they intended upon arrival, sometimes even staying permanently. We refer to this phenomenon as "unexpected staying". Had they correctly anticipated their actual duration of stay, unexpected stayers may well have changed how they lived: how much they integrated, saved, and invested in their country of destination. Unexpected staying might thus cause migrants to integrate less than their optimal level, which leads to lower performance on the destination country's labor market and lower social cohesion.

In this paper, we identify systematic causes for unexpected staying and suboptimal integration. To this end, we introduce general misperceptions about migrants' future utility and long-term wage prospects in a model with endogenous integration, savings, and long-term return decisions. We find that migrants *systematically* underestimate their probability of staying in the destination country if and only if their misperceptions involve pessimism about the destination country relative to the origin country. Using a long panel of migrants in Germany, we find empirical support for the relationship between pessimism at arrival, both in utility and wages, and long-term unexpected staying. In addition, we use our theoretical framework to identify specific biases or mislearnings which generate such general pessimistic misperceptions of utility or wage distributions. For instance, projection bias leads migrants to underestimate how much they will benefit from integrating and thus how likely they will stay. Another potential bias comes from migrants failing to realize that they observe non-representative or untrustworthy signals about wages in the destination country, which leads them to pessimistically misinfer their potential long-term labor market outcomes.

Other biases, such as present bias may generate unexpected staying for some migrant characteristics, but may also lead to unexpected leaving for others. We do not consider such mechanisms which only generate unexpected staying under specific circumstances. Also, we do not consider idiosyncratic shocks in migrants' lives, such as changes in marital status and having children, since a given migrant may be more likely to stay due to marrying, whereas another may be less likely to stay due to finding no partner. Individual-level shocks lead to unexpected staying *on average* only if migrants hold systematically biased beliefs over the distributions of such shocks.[1] Our approach

[1] Introducing such shocks into our model could be done by modelling uncertainty about the long-term level of

is thus complementary to research explaining the heterogeneity in staying behavior and identifying factors that correlate with the likelihood of staying longer than intended – such as specific life events (de Groot et al. (2011); Bettin et al. (2018)), high job satisfaction (Waldorf (1995)), age and feeling disadvantaged (van den Berg and Weynandt (2013)), as well as socio-economic and political conditions in the home country (Kirdar (2013)). From this perspective, this study is the first to identify pessimism about the destination country as a general and systematic cause for migrants to stay unexpectedly.

While understanding systematic mispredictions in migration duration is interesting *per se*, we are also interested in the impact it has on the host society, in particular by affecting the integration decisions. The literature links immigrants' intended duration of stay to various choices in the early phase of migration, such as investment in language skills (Dustmann (1999)) and savings' behavior (Sinning (2011))[2], with migrants' language skills being particularly important for their integration in the labor market (Gould et al. (1983); Dustmann and Soest (2002)). Understanding the causes of unexpected staying is thus also relevant for public policy.

In Section 3, we introduce the theoretical framework, which is applicable to migrants who participate in the labor market and can freely choose whether to return. It incorporates the essential features of the migrant's problem, namely uncertainty about long-term wages, endogenous integration and savings in the short term, and decision to return or to stay in the long term. Short-term and long-term decisions are intertwined: the more a migrant integrates or the less they save, the more likely they are to stay ex post. Conversely, the more a migrant anticipates to stay, the more they integrate and the less they save. These decisions are thus affected by the migrant's (mis)perceptions ex ante about their *core parameters*, i.e. their short-term and long-term preferences and their beliefs about labor market prospects. We first study multiple misperceptions of the core parameters and identify a condition about expected lifetime utility which implies that migrants integrate too little. This condition however does not ensure unexpected staying, which leads us to separately study misperceptions of each of these three core parameters in the three subsequent sections. For each individual misperception, we establish a necessary and sufficient condition for unexpected staying.

In Section 4, we first show that pessimism about long-term wages (in the form of first order stochastic dominance by the actual wage distribution) is necessary and sufficient for unexpected staying and sufficient for low integration. We then show that this type of pessimism can stem from prospective migrants who misinfer about the destination country from returning migrants in two ways. First, returning migrants are likely to have experienced worse than average outcomes in the destination country, and report this truthfully. Alternatively, in order to reduce the financial pressure induced by remittances sent to relatives living in the origin country, migrants may strategically underreport their income. In both cases, prospective migrants who (i) have incomplete information

integration.

[2]See Section 2 on related literature.

and (ii) fail to account for the selection or strategic concerns arising from the signals they observe will develop overly pessimistic expectations.

In Section 5, we consider misperceptions about utility in the long term. To generate unexpected staying, such misperceptions must involve an overestimation of the utility gap between the origin country (in case of return) and the destination country (in case of stay). We also show that a stronger form of pessimism, about returns to integration, implies low integration. In the next subsection, we show that projection bias and narrow bracketing lead to pessimistic misperceptions about long-term utility. Projection-biased migrants (Loewenstein et al. (2003)) underestimate how much they will adapt to the destination country, for instance due to thinking that they will always suffer from not being fluent in the local language, from not enjoying the food or the weather. Migrants who narrowly bracket their integration decision from their return decision may integrate in order to improve their day-to-day life, but fail to take into account how much more likely they are to stay as they become cumulatively more integrated.[3]

In Section 6, we consider misperceptions about utility in the short term. For unexpected staying to occur because of such misperceptions, migrants must overestimate the cost of integration efforts. Such migrants think ex ante that they will integrate less and save more than they actually will, so they overestimate their probability of return. The same holds if migrants underestimate the opportunity cost of savings. We show that present bias leads to underestimation of the cost of savings, but also to an *underestimation* of the cost of integration. Present-biased migrants may thus think that they will both save and integrate more than they actually will, which produces an ambiguous effect on the predicted probability of return. Present-biased migrants may or may not stay unexpectedly, depending on their characteristics.

In Section 7, we empirically investigate the link suggested by our theory between unexpected staying and migrants' initial pessimism about utility and about wage prospects. Using the German Socio-Economic Panel (GSOEP), we analyze a sample of migrants who arrived in Germany between 1982 and 2010. Among these migrants (excluding refugees), we consider as "unexpected stayers" those migrants who initially intended to stay temporarily and who are still in Germany as of 2017. We then study whether pessimism is associated with unexpected staying and with low integration and high savings. We measure pessimism about utility by exploiting information on migrants' life satisfaction. Specifically, we define pessimism about utility by comparing migrants' predicted life satisfaction in five years to their actual life satisfaction captured five years later. Pessimism about wage prospects is built on information about migrants' beliefs about wage increases in the future, compared to their actual wage growth two years later. Controlling for migrant socio-demographic characteristics, while we find that all migrants are overoptimistic about their life satisfaction, unexpected stayers are more pessimistic than other migrants. This lower level of optimism may be

[3] Note that while misinference (from Section 4) and projection bias both lead to too little integration and too much savings, narrow bracketing of integration has an ambiguous effect.

due to the mispredicted benefits of integration, which would be in line, for instance, with projection bias. We also find a significant association between pessimism about future wages and unexpected staying.

Our approach to identify the causes of a given behavior (here unexpected staying by migrants) consists in first introducing general misperceptions about utility and probabilities, and to derive the conditions under which these mispredictions lead to that behavior. This approach differs from the usual study of separate behavioral models designed for specific biases. Misperceptions are close to a sufficient statistic for unexpected staying, in so far as pessimistic misperceptions, no matter their cause, will unambiguously lead to more unexpected staying. We can thus study multiple biases that systematically lead to pessimistic misperceptions at once, which is inspired by a recent push to integrate robust insights from behavioral economics into applied economics (Mullainathan et al. (2012); Eliaz and Spiegler (2015); Chetty (2015); Handel and Schwartzstein (2018)). While our focus is on behavioral aspects, our approach equally applies to misperceptions that may be due to institutional or social factors such as biased reporting, information campaigns or aggregation in networks – always under the assumption that migrants fail to realize the full extent of the bias this entails.

Summing up, our theoretical framework identifies pessimism about the destination country as the central source of mispredicted duration of stay and suboptimal integration. Our empirical exercise provides supports for this finding. We also show through our theoretical framework that pessimism about future utility can be caused by projection bias and narrow bracketing, with projection bias also systematically lowering integration.

2 Related literature

Since life is full of surprises, the inaccuracy of return intentions per se is not puzzling, as it makes sense that some migrants adapt their plans in one way or another. Migrants are likely to reassess their plans through key life events in the destination country, and these events have different impacts depending on the stage of the migration process considered (Kley and Mulder (2010); Kley (2017)).

The systematic underestimation of migrants' duration of stay however suggests that more than mere uncertainty and luck are at play, and there is ample evidence for it. For instance, Schoorl (2011) finds that, while most economic migrants intend to stay temporarily in the Netherlands, only 40 percent of Turkish and 30 percent of the Moroccan migrants returned to their country of origin. This gap is significantly less pronounced for Italian and Spanish migrants, who return to their home country in larger proportions, and are more aligned with their initial intentions. Furthermore, Steiner and Velling (1994); van Baalen and Muller (2008) note that migrants' intended duration of stay keeps growing with the number of years spent in the destination country.[4]

[4] Ward (2017) is one case finding higher return rates than suggested by intentions, for migrants to the United States

While the empirical literature has provided explanations to unrealized return migration, these explanations do not seem to capture its systematic pattern. For instance, Waldorf (1995) finds that migrants who are satisfied with their job and residence are more likely to stay despite their initial intention to return. Lu (1999) documents that age and being a homeowner are positively associated with inconsistent intentions to move. Coulter (2013) shows that age, and changing levels of ties and commitments over the life course, explain the non-realization of past desires of residential mobility in the UK. Individuals' inability to realize their intentions to migrate might also be linked to a poor health condition (van Dalen and Henkens (2013)). van den Berg and Weynandt (2013) find that age and the feeling of being disadvantaged because of one's origins contribute to explain the gap between return intentions and actual stay. Hooijen et al. (2020) argue that recent university graduates are less likely to realize their intention to leave the region of study if they accumulated location-specific capital.

The systematic misprediction of migration duration has, to the best of our knowledge, not yet been analyzed in a general behavioral framework allowing to study the impact of incorrect beliefs and behavioral biases on individual decision-making (Odermatt and Stutzer (2019); Pinger et al. (2017); Dohmen et al. (2009)). Incorrect expectations have been shown to affect individuals' behavior in other life situations. The labor market, which tends to exhibit an excess of optimism, constitutes a relevant example. Spinnewijn (2015) shows that the unemployed tend to overestimate the speed at which they will find a new job, which results in insufficient search and savings. Krueger and Mueller (2016) show that unemployment duration has a very limited impact on workers' reservation wages in the US. Excess of optimism also applies to individuals holding a job, as Hoffman and Burks (2020) show that truck drivers over-estimate the number of miles they will run over the week, and they fail to update these estimations through the course of the week.

This excess of optimism about labour market prospects also appears to apply to immigrants (Borjas and Bratsberg (1996)). Shrestha (2020) finds that migrants without prior migration experience overestimate what they will earn on average by 26%, whereas in the case of Bangladesh this overestimation exceeds 50% (Bossavie et al. (2020)). In contrast, people who have not migrated (yet) often underestimate potential earnings from migrating (McKenzie et al. (2013); Seshan and Zubrickas (2017)). This can be due to strategic misrepresentation by previous migrants who understate their incomes to reduce pressure to share it with relatives (De Weerdt et al. (2019); Baseler (2020)).

Beyond improving our understanding of migrants' mispredicted length of stay, our paper is also linked to the important literature studying the link between immigrants' duration of stay and their integration in the destination country. It is indeed well-known that return intentions impact migrants' decision in many domains, such as integration and language acquisition (Dustmann (1999); Van Tubergen and Kalmijn (2009); Geurts and Lubbers (2017)), savings (Sinning (2011)),

in the late 19th and early 20th century.

remittances (Dustmann and Mestres (2010a); Delpierre and Verheyden (2014)), asset holdings in the origin and destination countries (Dustmann and Mestres (2010b), Chabé-Ferret et al. (2018)), and entrepreneurial investments in the home country (Ammassari (2004); Akwasi Agyeman and Fernández Garcia (2016)).

Another strand of literature highlights the benefits of integration for immigrants. First, integration directly improves the migrant's wellbeing through language proficiency (Amit and Bar-Lev (2014)), sense of belonging and perceived identity (Amit (2010)) and reduction in cultural distance (Angelini et al. (2015)). Importantly, the positive impact of migrant integration on their subjective well-being holds after controlling for individual socio-demographic characteristics, labor market status, as well as regional macroeconomic variables. Kogan et al. (2018) finds that immigrants who obtained naturalization, often seen as the most advanced level of integration, tend to report higher life satisfaction. Second, integration has an indirect effect on well-being through its benefits in terms of labor market performance. Knowledge of the local language in particular improves immigrants' labor market integration in many countries, including the US (Gould et al. (1983); Chiswick and Miller (2012)), the UK (Dustmann and Fabbri (2003)), France (Lochmann et al. (2019)), Australia (Guven and Islam (2015)) and Germany (Dustmann and Soest (2002)).

Our framework allows to address the socio-economic consequences of the systematic gap between intended and actual duration of stay. In particular, we study how immigrants' levels of ex-post language skills and savings relate to their optimal level had they not mispredicted their duration of stay.

3 A general framework for migrant misperceptions

We model a migrant who has recently arrived in the destination country. In the short run, they work and decide how much to integrate and save, and in the long run they decide whether to return to the origin country or to stay after learning new information. Our primary goal is to explore the circumstances under which a migrant *mispredicts* their return decision - with a secondary goal of exploring the implications this has for their migration and integration decisions. The framework is similar to that in Adda et al. (2019) who also explore factors driving temporary migration. They focus however on heterogeneity in location preferences and shocks to these preferences, and how this affects migrants' decisions to acquire human capital and return decision.

Given our goals, we have three periods in our model. In period 0, the migrant has just arrived in the destination country and predicts their future behavior, both short-term (savings and integration) and long-term (return) actions, but makes no decision and takes no action. Just as in Section 7, where we empirically analyze migrants' intentions to stay in the destination country, period 0 is the prediction period, during which the migrant has not received all the information yet, nor the ability to commit to actions, and thus may mispredict both of these. Then in period 1 (the "short term"), the migrant decides how much to integrate and save. Finally, in period 2 (the "long term") they learn their true utility from staying, including the economic situation in the destination country – modeled as the wage they will earn if they stay. Based on this, they decide whether to stay or to return. The migrant's location is denoted by L, with $L \in \{d, o\}$, where d and o denote destination and origin country, respectively.

In periods 1 and 2, migrants derive utility $u(c, I)$ from their consumption c and from the level of integration in the country in which they live, I. Integration increases direct utility ($\partial u(c, I)/\partial I > 0$) notably through the ability to understand the local language and increased quality of social interactions, the sense of belonging and perceived identity, and the reduction in cultural distance.[5] Throughout the paper, we will assume that the preferences $u(c, I)$ are separable ($\partial^2 u(c, I)/\partial I \partial c = 0$).[6]

Both levels of consumption c and I vary across periods. While the level of integration at arrival in the destination country I_1 is fixed, the period-2 level in the destination country I^d, can be increased thanks to integration efforts i made in period 1: $I^d = I_1 + i$.[7] These efforts are costly and reduce period-1 utility by $k(i)$, with $k'(i) > 0$, $k''(i) \geq 0$. In period 1, migrants also save money s into period-2 consumption with zero interest rate, so their period-1 consumption is noted $c_1(s) = s_0 + w_1 - s$, where s_0 are exogenous savings accumulated before period 1 and w_1 is the migrant's known wage in the destination country in period 1. Utility in period 1 is noted

$$v_1(s, i) \equiv u(c_1(s), I_1) - k(i),$$

where savings and integration efforts have convex costs in period 1: $\partial v_1/\partial s = -\partial u/\partial c < 0$, $\partial v_1/\partial i = -k'(i) < 0$, and $\partial^2 v_1/\partial s^2 = \partial^2 u/\partial c^2 \leq 0$, $\partial^2 v_1/\partial i^2 = -k''(i) \leq 0$.

In the second period, consumption and integration levels depend on the migrant's location. As mentioned above, the integration level in the destination country in period 2 is $I^d = I_1 + i$, whereas the integration level in the origin country is noted I^o. We assume that a migrant cannot be more integrated in the destination country than in the origin country, i.e. $I^d \leq I^o$.[8,9] Thus

[5] In the literature review, we have also mentioned that, in addition to the direct benefit of integration on migrant wellbeing, integration also improves migrants' labor market outcomes. In our model, this can be represented by an improvement of the distribution of long-term wages. Since this addition complicates the exposition without adding relevant insights, we restrict it to the Appendix, where we show that its introduction reinforces our results.

[6] While some of our results hold for any $u(c, I)$, others only hold if consumption and integration are weak complements, and others require full separability. Hence for ease of exposition, we choose to assume separability throughout.

[7] In order to simplify notations, we drop time subscripts whenever they are not needed. For instance, the fact that I^L, w^L and c^L have location superscripts implies that both locations are possible, which can only take place in period 2. Notations with superscripts thus only pertain to period 2. In contrast, I_1 can only denote the migrant's integration level in the destination country since it pertains to period 1, adding a superscript is thus not necessary.

[8] A sufficient condition for this is $k'(I^o - I_1) \to \infty$.

[9] It would make sense to assume that the migrant might become less well integrated in their home country, and thus underestimate how much less they will enjoy going back. This would have exactly the same effect as underestimating how much more they will enjoy staying, which is one of the biases we study.

integration benefits the migrant in case of stay ($\partial I^d/\partial i = 1 > 0 = \partial I^o/\partial i$).[10] Consumption in location L depends on the period-2 wage in location L, noted w^L. All wages are known and fixed with the exception of the period-2 wage in the destination country, w^d, which is uncertain and has distribution $F\left(w^d\right)$, and is an important factor determining the return decision. Given period-1 savings s, consumption in period 2 is thus either

$$\begin{aligned} c^d(s; w^d) &= w^d + s \text{ in case of stay,} \\ c^o(s) &= x \cdot (w^o + s) \text{ in case of return,} \end{aligned}$$

where $x \geq 1$ is the real exchange rate which accounts for the higher purchasing power in the origin country compared to the destination country ($\partial c^o/\partial s = x \geq 1 = \partial c^d/\partial s$). The value of returning to the origin country thus only depends on the level of savings, so we can write it as $v^o(s)$, whereas the value of staying in the destination country, $v^d\left(s, i; w^d\right)$, depends on period-1 savings s and integration efforts i, as well as on the realized wage w^d:

$$\begin{aligned} v^d(s, i; w^d) &\equiv u\left(c^d\left(s; w^d\right), I^d(i)\right), \\ v^o(s) &\equiv u\left(c^o(s), I^o\right). \end{aligned}$$

Having introduced the basic elements of the model, let us now study the migrant's decision problem, about integration and savings in period 1, and return migration in period 2, highlighting all the factors on which these decisions depend. Later on we introduce migrants' misperceptions about these elements and describe mechanisms leading to mispredictions about these decisions. First, we analyse the location choice, taking integration and savings as given.

Location decision: the reservation wage in the destination country In period 2, the migrant will return to the origin country if their utility in the destination country $v^d\left(s, i; w^d\right)$ is lower than their utility in the origin country $v^o(s)$. Return thus takes place if the realized wage in period 2 in the destination country is too low, that is if $w^d < w_R$, where the reservation wage w_R makes the migrant indifferent between both locations:

$$w_R\left(s, i; v^o(\cdot), v^d(\cdot)\right) \text{ is such that } v^d(s, i; w_R) = v^o(s). \tag{3.1}$$

Note that the indifference wage is uniquely determined by the migrant's period-1 actions (s, i) and their period-2 preferences $(v^o(\cdot), v^d(\cdot))$, but does not directly depend on period-1 preferences (v_1), or on $F\left(w^d\right)$. The threshold w_R together with the realized wage determines the migrant's decision to return or stay, and expectations over these two quantities determine expected return decisions. Thus a migrant who in period 0 perceives their long-term preferences to be $v^o(\cdot), v^d(\cdot)$ and who

[10] One could generalize this by allowing integration to also benefit the migrant in case of return. This, however, does not alter our results as long as integration is more beneficial in case of stay.

predicts period-1 actions as (s,i), thinks that they will return to the origin country with probability $F\left(w_R\left(s,i;v^o,v^d\right)\right)$. Any misprediction in short-term or long-term preferences, or in expected wages will thus generate a misprediction about the probability of return. Let us now turn to the choice of savings and integration in period 1.

Savings and integration decisions The choices of s and i are each based on a tradeoff between incurring a cost in period 1 in order to enjoy a higher expected utility in period 2. In period 1, the migrant anticipates that the (s,i) that they choose will lead to the period-2 reservation wage $w_R\left(s,i;v^o\left(\cdot\right),v^d\left(\cdot\right)\right)$, which also affects expected utility in period 2, noted $Ev_2\left(s,i;v^d,v^o,F\right)$, with

$$Ev_2\left(s,i;v^d,v^o,F\right)=\int_0^{w_R\left(s,i;v^o,v^d\right)}v^o(s)f(w^d)dw^d+\int_{w_R\left(s,i;v^o,v^d\right)}^{\infty}v^d(s,i)f(w^d)dw^d.^{11} \qquad (3.2)$$

The choice of (s,i) directly depends on the future benefits that the migrant expects to receive in period 2 through v^o or v^d, but also on the likelihood of returning, which depends on the wage distribution $F\left(\cdot\right)$ and on the reservation wage $w_R=w_R(s,i,v^o,v^d)$. One can thus already conclude that any misperception related to period-2 preferences or wage distribution will alter the choice of (s,i). Formally, the migrant's programme is to maximize $EV\left(s,i;v_1,v^d,v^o,F\right)$, with

$$EV\left(s,i;v_1,v^d,v^o,F\right)=v_1(s,i)+\delta Ev_2\left(s,i;v^d,v^o,F\right).$$

The optimal period-1 actions (s^*,i^*) are given by:

$$(s^*,i^*)=\arg\max_{s,i}EV\left(s,i;v_1,v^d,v^o,F\right).^{12} \qquad (3.3)$$

For these optimal actions, their period-2 reservation wage is $w_R\left(s^*,i^*;v^o,v^d\right)$, which results in the optimal probability to return to the origin country

$$p^*=F\left(w_R\left(s^*\left(v_1,v^d,v^o,F\right),i^*\left(v_1,v^d,v^o,F\right);v^o\left(\cdot\right),v^d\left(\cdot\right)\right)\right).$$

Let us now consider period 0, at which point migrants have beliefs, or perceptions, about $\left(v_1,v^d,v^o,F\right)$. Clearly, if these perceptions are correct, the migrant will correctly predict in period 0 that they will make these optimal choices. We now study the implications of potential misperceptions on predicted and actual short-term actions, and on the gap between predicted and actual probabilities of return migration in the long term.

[11] Note that, in addition to v^d and v^o, Ev_2 also depends on the migrant's perceived wage distribution F.

[12] We assume that there is a unique interior maximand for our maximization problem, which is ensured by the single-crossing of the curves representing the first order conditions (see Figure 1 in the Appendix A.1).

Introducing misperceptions about the migrant's core parameters We now introduce possible period-0 misperceptions by the migrant. The *"misperceived" core parameters* $\left(\widetilde{v}_1, \widetilde{v}^d, \widetilde{v}^o, \tilde{F}\right)$ are the channels through which a migrant may *unexpectedly stay.* In this section, we present a general class of misperceptions (possibly of all core parameters simultaneously) which leads migrants to opt for integration below its optimal level and savings above its optimal level. Since such a general misperception does not ensure the occurrence of unexpected staying, we devote the subsequent sections to the study of each individual misperception, and derive necessary and sufficient conditions which these individual misperceptions must satisfy in order to generate unexpected staying.

From here on, we assume that the migrant in period 0 predicts their future actions based on maximizing with respect to the misperceived core parameters $\left(\widetilde{v}_1, \widetilde{v}^d, \widetilde{v}^o, \tilde{F}\right)$ which differ in at least one element from the actual core parameters $\left(v_1, v^o, v^d, F\right)$. Again, the first decision to be discussed is the location choice, taking (s, i) as given. The migrant anticipates their location choice based on the misperceived reservation wage $\widetilde{w}_R = w_R\left(s, i; \widetilde{v}^d, \widetilde{v}^o\right)$, which is such that $\widetilde{v}^d\left(s, i; \widetilde{w}_R\right) = \widetilde{v}^o\left(s\right)$. In turn, the prediction made in period 0 about period-1 actions is

$$(\widetilde{s}, \widetilde{i}) = \arg\max_{s,i} E\widetilde{V}, \tag{3.4}$$

where

$$E\widetilde{V} = EV\left(s, i; \widetilde{v}_1, \widetilde{v}^o, \widetilde{v}^d, \tilde{F}\right).$$

These predicted actions lead to a prediction in period 0 of the probability of return migration noted

$$\widetilde{p} = \tilde{F}\left(w_R\left(\widetilde{s}, \widetilde{i}; \widetilde{v}^o, \widetilde{v}^d\right)\right), \tag{3.5}$$

where the mispredicted reservation wage results from mispredictions in both actions $(\widetilde{s}, \widetilde{i})$ and preferences $\left(\widetilde{v}^o, \widetilde{v}^d\right)$. It is crucial to note that this predicted probability $\widetilde{p}$ differs from the actual probability of return (ex post), which is

$$p_2 = F\left(w^R\left(\widehat{s}, \widehat{i}; v^o, v^d\right)\right), \tag{3.6}$$

where $\left(\widehat{s}, \widehat{i}\right)$ are the *actual* period-1 actions once utility v_1 is observed:

$$\left(\widehat{s}, \widehat{i}\right) = \arg\max_{s,i} E\widehat{V}, \tag{3.7}$$

where

$$E\widehat{V} = EV\left(s, i; v_1, \widetilde{v}^o, \widetilde{v}^d, \tilde{F}\right)$$

Note that if $\widetilde{v}_1$ does not differ from v_1, then the migrant's information set is unchanged when moving from period 0 to period 1 ($E\widehat{V} = E\widetilde{V}$). In other words, if $\widetilde{v}_1 = v_1$, the prediction made in period 0 about period-1 actions is correct ($(\widetilde{s}, \widetilde{i}) = (\widehat{s}, \widehat{i})$). This implies that only an incorrect perception about v_1 can generate a *misprediction of the migrant's period-1 actions.*[13,14]

[13] This does not mean that v_1 represents the "real" preferences, but represents the preferences with respect to which the actual choice is made. See our discussion of present bias.

[14] Had we ignored misperceptions about short-term utility, we could have specified the model without a period 0.

Comparing the return probability predicted in period 0, $\widetilde{p}$, to the actual return probability ex post, p_2, we get:

$$\widetilde{p} = \tilde{F}\left(w_R\left(\widetilde{s}, \widetilde{i}; \widetilde{v}^o, \widetilde{v}^d\right)\right) \lesseqgtr F\left(w^R\left(\widehat{s}, \widehat{i}; v^o, v^d\right)\right) = p_2.$$

Thus $\widetilde{p}$ can differ from p_2 because of differences between $(\tilde{s}, \tilde{i})$ and $(\hat{s}, \hat{i})$, which arise from misperceiving v_1, or between perceived distribution of wages $(\tilde{F})$ and the actual distribution (F), or between period-2 predicted utility $(\widetilde{v}^o, \widetilde{v}^d)$ and actual utility (v^o, v^d) – in short, from misperceptions in any of the core parameters. Our focus is on understanding which mispredictions lead to "unexpected staying", that is *over*estimating the probability of returning.

Definition 1. Unexpected staying *occurs if the migrant overestimates their probability of return migration:* $\tilde{p} > p_2$.

The notion of pessimism about the destination country is present throughout the discussion of migrant misperceptions leading to unexpected staying. While we do not derive clear conditions about combinations of misperceptions about $\left(\widetilde{v}_1, \widetilde{v}^d, \widetilde{v}^o, \widetilde{F}\right)$ which lead to unexpected staying, we can identify conditions which lead a migrant to integrate too little and save too much compared to a migrant who correctly perceives these core parameters. We define such conditions as general pessimism.

Definition 2. *A migrant is pessimistic about the destination country in general (**G-Pessimism**) if they underestimate the total marginal expected utility of integration and overestimate the total marginal expected utility of savings:* $\frac{\partial E\widetilde{V}(s,i)}{\partial i} \leq \frac{\partial EV(s,i)}{\partial i} \cap \frac{\partial E\widetilde{V}(s,i)}{\partial s} \geq \frac{\partial EV(s,i)}{\partial s}$ *for all* (s, i).

Overestimating the cost of integration is easily seen as a form of pessimism about the destination country. The reason we consider *under*estimating the cost of savings (the marginal utility of period-1 consumption) as pessimism is that lower savings make the destination country more desirable relative to the origin country, due to exchange rate effects. While general pessimism pertains to misperceptions about the total marginal utilities of integration and savings, these misperceptions will be broken down into misperceived short-term marginal costs and long-term marginal benefits of these actions (either in terms of distribution of outcomes or in terms of perceived utility).

The next point we need to introduce are migrant parameterizations z, which include the following migrant characteristics.

Definition 3. *A migrant parameterization* $z = (s_0, I_1, \delta, x, I^o, w^o)$ *is a vector of values that the migrant correctly perceives.*

In the following propositions, we will focus on conditions about misperceptions which are both sufficient and necessary to generate unexpected staying for all migrant parameterizations z.[15] This

[15] The essential characteristics that all these migrants should share for our theory to be applicable is that they are both willing and allowed to work it the destination country, and that they are able to decide their location in the long term.

implies that, conditional on a migrant's misperception satisfying these conditions, then they will unexpectedly stay, even if we know nothing about z – even if we do not know the specific migrant preferences or their starting point in the destination country. If the conditions did not generate for all z, then even with the same misperception there would be some migrants who would, under some circumstances, unexpectedly *leave* rather than unexpectedly stay. [16] By requiring that unexpected staying occurs for all z, we make stronger assumptions but weaken the data requirements for establishing whether the misperceptions will cause unexpected staying. These requirements are already high as we illustrate in Section 7. The reader should however keep in mind that we therefore do not cover all misperceptions that lead to unexpected staying, only those that lead to unexpected staying for any z. Having described the concepts of misperceptions, unexpected staying and migrant parameterizations, we can now state our first result.

Proposition 1. *Consider all possible migrants who share the same set of misperceptions* $\left(\widetilde{v}_1, \widetilde{v}^d, \widetilde{v}^o, \widetilde{F}\right) \neq \left(v_1, v^d, v^o, F\right)$. *Then,* $\widetilde{i} \leq i^*$ *and* $\widetilde{s} \geq s^*$ *for all* z *if their misperceptions* $\left(\widetilde{v}_1, \widetilde{v}^d, \widetilde{v}^o, \widetilde{F}\right)$ *satisfy G-pessimism.*

Proof. See Appendix. □

This result, which allows for a wide range of simultaneous misperceptions, states that if a migrant (of any characteristics) underestimates the net benefit of integration (comparing misperceptions of both its short-term costs and of its long-term benefits) and overestimates the net benefit of savings, then they are more likely to integrate less and to save more than they would if they had correctly perceived their core parameters. While this statement is fairly intuitive, it does not allow us to make clear statements about unexpected staying, as their predicted short-term actions in period 0 (which is what Proposition 1 is about) might differ from their actual actions in period 1, and it is those actions (together with their beliefs about the long term) which determine whether unexpected staying occurs.

To make clear statements about unexpected staying, we thus review individually each of the three types of misperceptions (about short-term utility v_1 and in particular the short-tem costs of integration and savings, about long-term utility $\left(v^d, v^o\right)$ and in particular the long-term benefits of these actions, and about the distribution of labor market outcomes F). In each of these sections, a proposition states conditions on misperceptions in order to generate unexpected staying for all z, followed by a proposition which describes how these mispredictions affect integration and savings in period 1.

[16] For instance, we will show that while misperceptions related to present bias can cause unexpected staying for some migrants, it can also generate unexpected leaving for others.

4 Mispredicting probabilities

In this section, we focus on a migrant with incorrect beliefs about *probabilities*, which means the perceived distribution of long-term wages in the destination country $\widetilde{F}(w^d)$ differs from the actual distribution $F(w^d)$. Other core parameters are correctly perceived: $\left(\widetilde{v}_1, \widetilde{v}^d, \widetilde{v}^o\right) = \left(v_1, v^d, v^o\right)$. In Subsection 4.1, we define which form of pessimism about wages is necessary and sufficient to lead to unexpected staying for all possible migrant preferences and characteristics. In Subsection 4.2, we discuss how migrants may develop such pessimistic wage expectations if they misinfer from other return migrants who either under-report their earnings or experienced particularly negative shocks.

4.1 Wage Pessimism and Unexpected Staying

Note that the reservation wage $w_R\left(s, i; v^d, v^o\right)$ is unaffected by beliefs about the wage distribution, thus the migrant correctly predicts it for any given (s, i). However, their predicted choice of $\left(\widetilde{s}, \widetilde{i}\right)$ is based on an incorrect wage distribution:

$$E\widetilde{V}\left(s, i; v_1, v^d, v^o, \widetilde{F}\right) = v_1(s, i) + \delta\left(\int_0^{w_R\left(s,i;v^d,v^o\right)} v^o(s)\tilde{f}(w^d)dw^d + \int_{w_R\left(s,i;v^d,v^o\right)}^{\infty} v^d(s, i)\tilde{f}(w^d)dw^d\right).$$

As a result, their period-1 predicted actions are not optimal: $\left(\widetilde{s}, \widetilde{i}\right)$ differs from (s^*, i^*), i.e. the choice of a migrant who correctly perceives the distribution of wages in period 2 (3.3). In addition, even fixing (s, i), they directly mispredict their probability of return $\widetilde{F}\left(w_R\left(\widetilde{s}, \widetilde{i}; v^d, v^o\right)\right)$ for all (s, i). Based on $\left(\widetilde{s}, \widetilde{i}\right)$, they predict their return probability to be $\widetilde{p} = \widetilde{F}\left(w_R\left(\widetilde{s}, \widetilde{i}; v^d, v^o\right)\right)$, whereas their actual probability of return is $p_2 = F\left(w_R\left(\widetilde{s}, \widetilde{i}; v^d, v^o\right)\right)$. Let us establish which misperception $\widetilde{F}$ is necessary and sufficient to generate unexpected staying. To this end, we introduce the following definition.

Definition 4. *A migrant is **pessimistic about probabilities** in the destination country (**P-pessimistic**) if the misperceived wage distribution $\tilde{F}$ is first-order stochastically dominated by the actual distribution F, i.e. $\tilde{F}(w^d) \geq F(w^d)$ for all w^d.*

A P-pessimistic migrant expects lower wages than they are likely to receive in the destination country.[17] We can now state the relationship between pessimism about the wage distribution and unexpected staying.

Proposition 2. *Consider all possible migrants with the same misperception $\widetilde{F}$. Unexpected staying occurs for all $\left(z, v_1, v^d, v^o\right)$ if and only if their misperception $\tilde{F}$ satisfies P-pessimism.*

[17] The migrant could also be relatively pessimistic about the destination country because of an excess of optimism about long-term wages in the origin country. For the sake of simplicity, we only model misperceptions about the destination country's wage distribution, but this alternative would yield the same results.

Proof. See Appendix. □

By all possible migrants with the same misperception $\widetilde{F}$, we mean all migrants with (i) any possible values of (correctly perceived) core parameters $\left(v_1, v^d, v^o\right)$ and any possible $z = (s_0, I_1, \delta, x, I^o, w^o)$, and (ii) for whom the only common feature is their misprediction of F. While the fact that P-pessimism is a sufficient condition to generate unexpected staying is relatively obvious, let us explain why it is necessary. Consider a misperception $\widetilde{F}$ which does not satisfy P-pessimism and which generates unexpected staying for some particular migrant. Since this condition does not satisfy P-pessimism, we can claim that there exists a migrant with different characteristics for whom $\widetilde{F}$ leads to unexpected *leaving*.

The intuition of this proposition is that P-pessimism makes the migrant think that they are more likely to obtain a wage below the threshold $w_R\left(\widetilde{s}, \widetilde{i}; v^d, v^o\right)$ than they actually are, which leads them to overestimate their probability of return. Let us now discuss how the actual short-term actions $\left(\widehat{s}, \widehat{i}\right)$ of a P-pessimistic migrant (which in this case are identical to predicted actions $\left(\widetilde{s}, \widetilde{i}\right)$) differ from the actions of a migrant whose perceptions are fully correct (s^*, i^*).

Proposition 3. *Consider all possible migrants with the same misperceptions* $\widetilde{F}$*. Then,* $\widetilde{s} = \widehat{s} > s^*$ *and* $\widetilde{i} = \widehat{i} < i^*$ *for all* $\left(z, v_1, v^d, v^o\right)$ *if their misperception* $\widetilde{F}$ *satisfies P-pessimism.*

Proof. See Appendix . □

First note that this result is a corollary of Proposition 1. Also note that by ruling out short-term misperceptions about utility, predicted and actual period-1 actions are identical ($\widetilde{s} = \widehat{s}$ and $\widetilde{i} = \widehat{i}$). The P-pessimistic migrant saves too much because they overestimate the expected returns to savings, which have higher utility in the origin country. Similarly, they integrate too little because they underestimate the expected returns to integration. Because of this, they are more likely to return ex post than if they had saved and integrated optimally: $F\left(w_R\left(\widehat{s}, \widehat{i}; v^d, v^o\right)\right) > F\left(w_R\left(s^*, i^*; v^d, v^o\right)\right)$.

Let us now discuss specific situations which might lead to P-pessimism.

4.2 Misinference from migrants as cause for wage pessimism

As described in the related literature section, there is mixed evidence on the degree of pessimism of people living in migrant-sending countries about job prospects in destination countries.[18]

On the one hand, McKenzie et al. (2013); Seshan and Zubrickas (2017) have documented that potential earnings from migration tend to be underestimated in the origin country. On the other hand, Shrestha (2020) finds that potential work migrants from Nepal to Malaysia and the Persian Gulf countries overestimate earnings in these destination countries by 26% on average. Bossavie et al. (2020) show that such overestimation exceeds 50% for migrants from Bangladesh.

[18] In contrast, non-migrants living in developed countries seem systematically optimistic (Spinnewijn (2015), Krueger and Mueller (2016), Hoffman and Burks (2020)).

When forming expectations about their own prospects, prospective migrants may base these expectations on the labor market outcomes in the destination country reported by return migrants. However, the outcomes of this subgroup are likely to be negatively selected among the whole population of migrants. Potential migrants in the origin country may fail to account for this negative selection, instead treating the sample of return migrants as a representative sample for outcomes in the destination country. In a lab experiment, Enke (2020) finds that participants interpret selected signals as if they were an unbiased sample of signals, in a setting where inference is simpler and more transparent than in the case of return migrants. Additionally, migrants in the destination country may have incentives to underreport their income, in order to reduce pressure to redistribute it to relatives (De Weerdt et al. (2019); Baseler (2020)). Various models have been proposed in which agents misinfer in such strategic settings, along with evidence of neglect of such strategic concerns and selection issues (Eyster and Rabin (2005); Esponda (2008)).

5 Mispredicting long-term preferences

In this section, we consider errors that are exclusively due to the migrant misperceiving the utility of outcomes in the long term (i.e. in period 2). While they perceive preferences to be $v^d(\cdot,\cdot;\cdot)$ and $v^o(\cdot)$ ex post, their prediction of these preferences in periods 0 and 1 are $\widetilde{v}^d(\cdot,\cdot;\cdot) \equiv \widetilde{u}^d\left(c^d, I^d\right)$ and $\widetilde{v}^o(\cdot) \equiv \widetilde{u}^o(c^o, I^o)$. The first part shows that pessimism about long-term preferences is necessary and sufficient for unexpected staying under all circumstances, while the second part highlights several mechanisms that lead to such pessimism.

5.1 Long-term Pessimism and Unexpected Staying

Unlike mispredicted probabilities, misperceptions about long-term preferences have implications on $\tilde{w}_R(s,i) = w_R\left(s,i;\widetilde{v}^d,\widetilde{v}^o\right)$, the misperceived wage threshold which makes the migrant *think* that they will be indifferent between both locations, that is $\widetilde{v}^d(s,i;\tilde{w}_R) = \widetilde{v}^o(s)$.

A migrant mispredicting their future utility chooses $(\widetilde{s},\widetilde{i}) = \arg\max_{s,i} E\widetilde{V}$, with $E\widetilde{V}$ given as follows:

$$E\widetilde{V} = v_1(s,i) + \delta\left(\int_0^{\tilde{w}_R(s,i)} \widetilde{v}^o(s) f(w^d)dw^d + \int_{\tilde{w}_R(s,i)}^{\infty} \widetilde{v}^d(s,i;w^d) f(w^d)dw^d\right).$$

Note that anticipated and actual savings and integration decisions are identical ($(\tilde{s},\tilde{\imath}) = (\hat{s},\hat{\imath})$) since migrants correctly predict short-term preferences. Given these savings and integration decisions, their perceived probability of staying is $\widetilde{p} = F\left(w_R\left(\tilde{s},\tilde{\imath};\widetilde{v}^d,\widetilde{v}^o\right)\right)$. Since ex post period-2 preferences differ from the initial perceptions, their actual probability of returning is in fact $p_2 = F\left(w_R\left(\tilde{s},\tilde{\imath};v^d,v^o\right)\right)$. *Unexpected staying thus takes place if* $\widetilde{p} > p_2$, or equivalently if $w_R\left(\tilde{s},\tilde{\imath};\widetilde{v}^d,\widetilde{v}^o\right) > w_R\left(\tilde{s},\tilde{\imath};v^d,v^o\right)$.

Let us now specify a pessimistic form of misperception about $(\widetilde{v}^o,\widetilde{v}^d)$.

Definition 5. *A migrant is pessimistic about long-term utility in the destination country (**LT-pessimistic**) if* $\widetilde{v}^d\left(s,i;w^d\right)-\widetilde{v}^o\left(s\right)\leq v^d\left(s,i;w^d\right)-v^o\left(s\right)$ *for any given* i*,* s*, and* w^d*.*

This type of pessimism is necessary and sufficient for unexpected staying, as formalized in the following proposition.

Proposition 4. *Consider all possible migrants with the same misperceptions* $\left(\widetilde{v}^d,\widetilde{v}^o\right)$*. Unexpected staying occurs for all* (z,v_1,F) *if and only if their misperception* $\left(\widetilde{v}^o,\widetilde{v}^d\right)$ *satisfies LT-pessimism.*

By all migrants with the same misperceptions $\left(\widetilde{v}^d,\widetilde{v}^o\right)$, we mean migrants with all possible values of (correctly perceived) core parameters (v_1,F) and characteristics $z=(\delta,x,s_0,w^o,I_1,I^o)$, for whom the only common feature is their misprediction of $\left(v^d,v^o\right)$. LT-pessimism being necessary for all possible migrants means that if a specific migrant stays unexpectedly even though their misperception does not satisfy LT-pessimism, then there always exists a migrant with the same misperception but different characteristics who instead *leaves* unexpectedly.[19] By making the destination country appear less appealing than it will be ex post, LT-pessimism makes the migrant think that in order to stay they will need a higher reservation wage than the actual (ex post) reservation wage: $w_R\left(\tilde{s},\tilde{\imath};\widetilde{v}^d,\widetilde{v}^o\right)\geq w_R\left(\tilde{s},\tilde{\imath};v^d,v^o\right)$. Consequently, as stated in Proposition 4, pessimism implies that the migrant overestimates their probability of return migration.

The second question is how a pessimistic migrant actually behaves in terms of savings and integration in period 1 $(\widehat{s},\widehat{\imath})$ compared to a migrant whose utility is correctly predicted (s^*,i^*). Before stating Proposition 5, let us introduce a stronger version of pessimism about long-term returns to savings and integration which implies LT-pessimism.

Definition 6. *Pessimism about long-term returns in the destination country **(LTR-pessimism)** is such that the migrant underestimates future marginal benefits of integration and savings in the destination country (*$\partial\widetilde{v}^d/\partial s\leq\partial v^d/\partial s$ *and* $\partial\widetilde{v}^d/\partial i\leq\partial v^d/\partial i$ *for all* s *and* i*).*

It is easy to see via integration that LTR-pessimism implies LT-pessimism.

Proposition 5. *Consider all possible migrants with the same misperceptions* $\left(\widetilde{v}^d,\widetilde{v}^o\right)$*. Then,* $\widetilde{s}=\widehat{s}>s^*$ *and* $\widetilde{\imath}=\widehat{\imath}<i^*$ *for all* (z,v_1,F) *if their misperception* $\widetilde{v}^d$ *satisfies LTR-pessimism.*

Proof. See Appendix. □

First note that this result is a corollary of Proposition 1, and that by ruling out short-term misperceptions, predicted and actual period-1 actions are identical ($\widetilde{s}=\widehat{s}$ and $\widetilde{\imath}=\widehat{\imath}$). Proposition 5 implies that due to higher savings and lower integration, LTR-pessimistic migrants are more likely to return ex post than if they had integrated and saved optimally: $F\left(w_R\left(\widehat{s},\widehat{\imath};v^d,v^o\right)\right)>F\left(w_R\left(s^*,i^*;v^d,v^o\right)\right)$. However, when they stay they are obviously worse off.

[19] A more detailed discussion of this necessary condition is provided in Section 4 (under Proposition 2).

5.2 Projection bias over adaptation causes long-term pessimism

Weynandt (2014) suggests that projection bias (Loewenstein et al. (2003)) may contribute to the gap between expected and realized return migrations in the German SOEP. Projection bias refers to the tendence of people to perceive their future tastes as more similar to their current tastes than they are.[20] Initially, migrants dislike the food, the weather and the customs, all of which are different from their origin country. The lack of speaking the local language in particular may make them feel alienated.

Let us apply the model of simple projection bias from Loewenstein et al. (2003) in our framework. Under projection bias, the migrant perceives the utility in the destination country in period 2 as lying between the actual utility it will yield in period 2 with I^d and the utility this would yield in period 1, with integration I_1. Therefore, for all $\left(s, i; w^d\right)$, the utility from consumption and integration in the destination country $u\left(c^d\left(s; w^d\right), I^d\left(i\right)\right) = v^d(s, i; w^d)$ is perceived lower than it will be once the migrant has settled, grown used to the customs, and learned the local language.[21]

Following Loewenstein et al. (2003), we formally get the following:

$$\tilde{u}(c^d, I^d | I_1, \alpha) = \alpha u(c^d, I_1) + (1 - \alpha) u(c^d, I^d),$$

where $\tilde{u}(c^d, I^d | I_1)$ denotes the period-2 utility as perceived in period 0, when integration is I_1, and $\alpha \in [0, 1]$ measures the degree of projection bias. When $\alpha = 0$, there is no bias and the migrant perceives utility accurately ; when $\alpha = 1$, the migrant is fully biased and perceives future utility to be identical to utility with an integration level I_1.

We can now directly check that projection bias leads to long-term pessimism, since misperceived utility is lower than actual utility, for all consumption and integration levels:

$$\tilde{u}^d(c^d, I^d | I_1, \alpha) = u^d(c^d, I^d) - \alpha \left(u^d(c^d, I^d) - u^d(c^d, I_1)\right) \leq u^d(c^d, I^d),$$

since for any given c^d, $\left(u(c^d, I^d) - u(c^d, I_1)\right) > 0$ by $\partial u / \partial I > 0$. Moreover, the migrant does not misperceive their home country utility, since they no longer increase their level of integration there (we can allow for losing integration which will merely strengthen the results). Therefore we have that $\widetilde{v}^d\left(s, i; w^d\right) - \widetilde{v}^o\left(s\right) \leq v^d\left(s, i; w^d\right) - v^o\left(s\right)$, which is the definition of LT-pessimism. Note that projection bias also satisfies strong LT-pessimism. Hence, applying Proposition 4 and 5, one can conclude that projection bias leads to unexpected staying, insufficient integration and too much savings.

[20] Evidence for projection bias due to short-term fluctuations includes drug addiction (Badger et al. (2007)), sexual arousal (Ariely and Loewenstein (2006)), and effort exertion (Augenblick and Rabin (2018); Bushong and Gagnon-Bartsch (2020)), as well as for people mispredicting their habit formation for gym attendance.

[21] Unlike Dustmann and Görlach (2016) who focus on the role of skill and human capital, we focus on how integration increases the direct utility of consumption irrespective of any productivity gains, which is the channel through which projection bias operates. We however present in Appendix A.3.2 an extension of the model in which integration can improve the distribution of long-term wages, and show that unexpected staying can also occur if migrants underestimate this additional effect.

5.3 Underestimating incidental integration and Pessimism

Narrow bracketing refers to situations in which people should account for the impact of one choice on all other choices, but instead neglect such impacts (Read et al. (1999)).[22] Let us now model a form of narrow bracketing that we call incidental integration, and which generates unexpected staying but does not systematically lead to below-the-optimum integration. Suppose that integration efforts can be divided into two categories, namely basic day-to-day interactions with natives j, and active efforts (i in our baseline model) which structurally increase the level of long-term integration, such as language trainings. While day-to-day efforts j made in period 1 are driven by interactions with natives (during which the person exerts effort in order to get consumption, such as ordering food in a shop, etc), they also incidentally increase the level of long-term integration.

$$I^d = I_0 + j + i.$$

Formally, the migrant has an additional decision variable in period 1, j, which we treat as being entirely fixed by the choice of consumption (and hence by the choice of savings).[23]

The key point here is that if the migrant narrowly brackets j from i, they ignore the effect of j when they predict their long-term integration level:

$$\widetilde{I}^d = I_0 + (1-\gamma)\, j + i \leq I^d.$$

Underestimating long-term integration directly implies that period-2 utility in the destination country is also underestimated:

$$\widetilde{v}^d(s, i; w^d, \gamma) = u^d\left(w^d + s, I_0 + (1-\gamma)\, j + i\right) \leq v^d(s, i; w^d).$$

Since a migrant has no day-to-day interactions in their origin country, they cannot underestimate the impact of incidental integration in the origin country, so that narrow bracketing leads to unexpected staying.[24] The migrant does not however underestimate the returns to integration as required for proposition 5, hence we cannot conclude that the person always integrates less and saves more.

[22] Evidence for this type of behavior goes from the failure of people (i) to combine multiple lotteries and thus losing out on the benefits of diversification (Kahneman and Tversky (1979); Redelmeier and Tversky (1992); Rabin and Weizsäcker (2009)); (ii) to combine multiple decisions for sharing money among other people (Exley and Kessler (2018); Ellis and Freeman (2020)); (iii) to combine additional effort with baseline effort (Fallucchi and Kaufmann (2020)).

[23] Alternatively we could model it directly as an additional decision variable, with utility in period 1

$$v_1(s, i, j) = u(c_1(s), I_1(j)) - k(i) - l(j).$$

[24] Instead, one example where v^o is *over*estimated (which magnifies this effect), is if integration in the origin country decays with time spent abroad and the migrant neglects this decay.

6 Mispredicting short-term preferences

We finally turn our focus to models in which in period 0 the migrant mispredicts their utility in period 1: they perceive v_1 to be $\tilde{v}_1$, and therefore mispredict their actions in period 1, while other core parameters $\left(v^d, v^o, F\right)$ are correctly perceived. In Subsection 6.1, we show how pessimism about the (utility) costs of short-term actions leads to unexpected staying. In Subsection 6.2, we illustrate this through a naive and present-biased migrant, who may think ex ante that they will integrate and save intensively, yet end up avoiding such efforts come the moment. We show that such bias does not satisfy pessimism about costs and thus does not lead to unexpected staying under all parameterizations.

6.1 Short-term Pessimism and Unexpected Staying

Consider migrants who misperceive v_1 as $\tilde{v}_1$. They perceive everything else correctly though, and thus predict their choices to be $(\tilde{s}, \tilde{\imath})$ based on maximizing:

$$\max_{s,i} \tilde{v}_1(s,i) + \delta E v_2\left(s, i; v^d, v^o, F\right),$$

where the correctly perceived period-2 expected utility $Ev_2\left(s, i; v^d, v^o, F\right)$ is defined in (3.2).

Their true choices in period 1 instead turn out to be $(\hat{s}, \hat{\imath})$ based on $v_1(\cdot,\cdot)$ rather than $\tilde{v}_1(\cdot,\cdot)$. This implies that in period 0, the migrant predicts that they will choose $(\tilde{s}, \tilde{\imath})$ in period 1 and that they will return with probability $\widetilde{p} = F\left(w_R\left(\tilde{s}, \tilde{\imath}; v^d, v^o\right)\right)$ in period 2, whereas one period later, they adapt their choices to $(\hat{s}, \hat{\imath})$ and have an actual probability to return $p_2 = F\left(w_R\left(\hat{s}, \hat{\imath}; v^d, v^o\right)\right)$.

Here the migrant can misperceive short-term preferences $\widetilde{v}_1(s,i) \equiv \widetilde{u}\left(c_1(s), I_1\right) - \widetilde{k}(i)$. Let us define the notion of pessimism we will need for short-term mispredictions:

Definition 7. *A migrant is pessimistic about the destination country in the short term (**ST-pessimism**) if the migrant*

- *overestimates the period-1 cost of integration: for any given i,*

$$\left|\partial\widetilde{v}_1/\partial i\right| = \tilde{k}'(i) \geq k'(i) = \left|\partial v_1/\partial i\right|,$$

- *underestimates the period-1 cost of savings: for any given s and w^d,*

$$\left|\partial\widetilde{v}_1/\partial s\right| = \partial\widetilde{u}^d\left(w^d + s, I_1\right)/\partial s \leq \partial u^d\left(w^d + s, I_1\right)/\partial s = \left|\partial v_1/\partial s\right|.$$

Overestimating the cost of integration is easily seen as a form of pessimism about the destination country. The reason we consider *under*estimating the cost of savings (the marginal utility of period-1 consumption) as pessimism is that higher savings make the destination country less desirable relative to the origin country due to differences in purchasing power.

Proposition 6. *Consider all possible migrants with the same misperception $\widetilde{v}_1$, which stems from a linear separable and strictly monotonic bias. Unexpected staying occurs for all $\left(z, v^d, v^o, F\right)$ if and only if their misperception $\widetilde{v}_1$ satisfies ST-pessimism.*

Proof. See Appendix. □

By all migrants with the same misperceptions $\widetilde{v}_1$, we mean migrants with all possible values of (correctly perceived) core parameters $\left(v^d, v^o, F\right)$ and all possible initial values $z = (\delta, x, s_0, w^o, I_1, I^o)$, for whom the only common feature is their misprediction of v_1. ST-pessimism being necessary means that no other short-term misperception is able to systematically generate unexpected staying for all these migrants.

To conclude this subsection, let us compare the ST-pessimistic migrant's predicted actions $(\tilde{s}, \tilde{\imath})$ to their actual actions $\left(\widehat{s}, \widehat{\imath}\right)$, which in this case are identical to the optimal actions of an unbiased migrant (s^*, i^*).

Proposition 7. *Consider all possible migrants with the same misperception $\widetilde{v}_1$. Then, $\tilde{s} \geq \widehat{s} = s^*$ and $\tilde{\imath} \leq \widehat{\imath} = i^*$ for all $\left(z, v^d, v^o, F\right)$ if their misperception $\widetilde{v}_1$ satisfies ST-pessimism.*

This result is again a corrolary of Proposition 1, in which actual short-term actions are identical to optimal actions ($\widehat{s} = s^*$ and $\widehat{\imath} = i^*$) since at period 1 the migrant correctly perceives the future (both in probability and in utility). Intuitively, if in period 0 the migrant overestimates the cost of actions which favor staying (integration) and underestimates the cost of actions which favor returning (savings), they will choose $\tilde{s} \geq \widehat{s}$ and $\tilde{\imath} \leq \widehat{\imath}$. These predicted actions will make them overestimate their probability of return in period 2 , which in turn implies unexpected staying. In period 1, they will realize that actually saving is more costly and that integration is less costly than they thought, and thus effectively save less, integrate more, and end up staying more frequently than they had predicted.

6.2 Present bias does not systematically cause unexpected staying

Proposition 6 applied to naive present bias does not systematically lead to unexpected staying. Note first that sophisticated time-inconsistent preferences cannot explain unexpected staying, since sophisticated migrants predict their own actions correctly, which can never lead to mispredictions of return intentions. We thus focus exclusively on the naive case. Taking the specific form from O'Donoghue and Rabin (1999), migrants have $\beta - \delta$ preferences in which the β parameter uniformly discounts all future periods. So in period 1, the migrant applies the following programme:

$$\left(\widehat{s}, \widehat{\imath}\right) = \arg\max_{s,i} g_1(s, i) + \beta\delta E v_2\left(s, i; v^d, v^o, F\right),$$

where $g_1(s, i)$ is the period-1 utility function in the presence of present bias. Equivalently, one can rewrite this programme by dividing the objective function by β to obtain $\left(\widehat{s}, \widehat{\imath}\right) = \arg\max_{s,i} EV$, where

$$EV = v_1(s,i) + \delta E v_2 \left(s, i; v^d, v^o, F\right),$$

where $v_1(s,i) = g_1(s,i)/\beta$. In period 0, however, the naive migrant misperceives β and overestimates how much they value the future in period 1. Instead of β, the migrant expects to weigh the future by $\widetilde{\beta} \geq \beta$ – to be more patient than they are. In period 0 they think that they will apply the following programme in period 1:

$$\left(\widetilde{s}, \widetilde{i}\right) = \arg\max_{s,i} g_1(s,i) + \widetilde{\beta}\delta E v_2 \left(s, i; v^d, v^o, F\right).$$

Dividing the former equation by $\widetilde{\beta}$, this programme is equivalent to our canonical programme with misperceptions $\left(\widetilde{s}, \widetilde{i}\right) = \arg\max_{s,i} E\widetilde{V}$ where:

$$E\widetilde{V} = \widetilde{v}_1(s,i) + \delta E v_2 \left(s, i; v^d, v^o, F\right).$$

and misperceived period-1 utility is

$$\widetilde{v}_1(s,i) = \frac{\beta}{\widetilde{\beta}} v_1(s,i).^{25}$$

Since $\frac{\beta}{\widetilde{\beta}} \leq 1$, we have that $\widetilde{v}_1(s,i) \leq v_1(s,i)$, which means that in period 0, the migrant underestimates how much they will value period-1 utility when they will make period-1 decisions. In this case, it is important to note that the migrant uniformly underestimates the costs of *both period-1 actions*, namely *savings* ($|\partial \widetilde{v}_1(s,i)/\partial s| = \frac{\beta}{\widetilde{\beta}} |\partial v_1(s,i)/\partial s| < |\partial v_1(s,i)/\partial s|$) *and integration* ($\widetilde{k}'(i) = \frac{\beta}{\widetilde{\beta}} k'(i) < k(i)$). This violates the ST-pessimism condition, hence there are parameterizations of the problem such that present bias implies unexpected leaving rather than unexpected staying.

7 Stylized Facts from German SOEP

In this section, we aim to document whether, in line with theory, initial pessimistic misperceptions (about utility or wage prospects) increase the likelihood that migrants end up staying in Germany beyond their initial intentions at arrival (i.e. the likelihood of staying unexpectedly). To this end, we exploit information on a sample of migrants from the German Socio-Economic Panel (GSOEP). This dataset permits to follow migrants over several decades while providing information on the multiple dimensions captured by our model: immigrants' initial intentions (to stay permanently or temporarily), proxies for their initial levels of pessimism, their short-term levels of integration and savings, and their long-term actual location.

[25] To see this, note that the actual period-1 utility is $v_1(s,i) = \frac{1}{\beta} z_1(s,i)$, and its misperceived counterpart is $\widetilde{v}_1(s,i) = \frac{1}{\widetilde{\beta}} z_1(s,i) = \frac{\beta}{\widetilde{\beta}} v_1(s,i)$.

The data and sample criteria The GSOEP is a survey which provides household- and individual-level data for a representative sample of the population in Germany on a yearly basis. This population includes immigrants observed since the first wave in 1984 to the last at our disposal in 2017. In addition to standard demographic and socio-economic characteristics, the panel offers valuable information about immigrants' country of birth, year of arrival and last known location (i.e. whether the migrant left Germany), in addition to the previously mentioned variables essential for this analysis. Our sample consists of migrants who had either temporary or permanent initial intentions, and who have either left Germany, or stayed. To ensure that pessimism, intentions, and integration and savings were captured in the short term (i.e. close to arrival in Germany) and that the actual location ex post is informative of whether initial intentions were correct or not in the long term, we apply the following sample selection criteria.[26] First, since we need initial return intentions, we exclude migrants whose first reply in the survey takes place at a time when they were already settled, had already made integration decisions and had acquired sufficient information about their career prospects. We therefore restrict our sample to migrants who were first interviewed at most two years after their arrival in Germany.[27] Second, we restrict the sample to migrants who arrived in Germany no later than 2010, in order to allow enough time for a possible return, with our last sample year being 2017. Third, to ensure that migrants' integration is a choice rather than a constraint, we focus on individuals aged at least 18 when they arrived in Germany.

Key variables and descriptive statistics Our focus is on the link between pessimism at arrival and unexpected staying: still being in Germany beyond the initially intended duration.

We proxy pessimism about utility by using the migrant's actual and predicted life satisfaction (LS). Upon arrival (at the same time as when they state their intention to be temporary or permanent), migrants formulate their anticipated life satisfaction in five years. We then exploit the actual life satisfaction that they state five years later to build our measure of pessimism. This pessimism is simply the gap between their actual life satisfaction ex post and the life satisfaction that they predicted they would have 5 years earlier, both variables being measured on a scale from 0 to 10. A migrant's $LS\_pessimism$ score thus theoretically ranges between -10 and 10, where 10 corresponds to the most pessimistic migrant, whose ex post life satisfaction is 10, whereas they had predicted that it would be 0.

The measure of pessimism about wages follows a similar logic, as it is based on the GSOEP question about the perceived probability of obtaining a pay raise in the following two years. From this question, whose response modalities changed in 1999, we create a binary indicator which

[26] In addition to the criteria listed below, we also drop immigrants who reside in East Germany as well as respondents who were late repatriates (i.e. "Spätaussiedler" : immigrants of German descent that lived in the Eastern block) or who have a refugee status. The immigration and return decisions of these migrant groups are likely affected by different institutional settings, migration motives and constraints.

[27] Therefore, the first migrants in our sample arrived in Germany in 1982 at the earliest.

takes the value 1 if the respondent expected a rise to occur with more than 50% probability or considered it to "definitely" or "probably" happen. Otherwise, the indicator takes the value 0.[28] Again, we focus on answers to this question which were provided within the first two years after arrival, and compare it to their actual labor income growth. We attribute a value of 1 to the actual labor income growth if labor income grew by at least 5% over two years, and 0 otherwise. A migrant is thus pessimistic ($wage\_pessimism = 1$) if they had not expected a pay raise ($expected\ pay\ raise\ indicator = 0$), whereas they did experience it ex post, whereas a migrant turns out to be optimistic ($wage\_pessimism = -1$) if they expected higher wages ($expected\ pay\ raise\ indicator = 1$), but did not ex post.

To be an unexpected stayer, a migrant needs to (i) have initial temporary intentions and (ii) to stay in Germany in the long term.[29] Our main measure of unexpected staying, "UnexpS" is a binary variable which takes value 1 if the migrant stays in Germany although they had expected to have left by 2017, and value 0 if they left Germany, or have correctly predicted to still be in Germany. We rely on this binary variable for several reasons. First, intended duration of stay expressed in years is only available for a subset of migrants (some do not reply to this question) who state a temporary migration intention. Hence, an arbitrary imputation, for instance based on an average life expectancy, would be needed to create this variable for respondents with a permanent intention, as they were not asked about their intended duration of stay. Second, the number of intended years of stay has a larger within-individual volatility than the binary measure of temporary/permanent intention (i.e. the number of intended years of stay can vary while the respondent keeps stating a constant temporary intention). Hence, values of this alternative variable would depend on the survey year selected to construct it, and we prefer to avoid this arbitrary choice.

We first present descriptive statistics of our sample in (1) and then split it between unexpected stayers (2) and those who either have left Germany, or have correctly predicted to remain in Germany until at least 2017 (3). Column (4) provides t-tests on the mean differences between subsamples (2) and (3). Table 1 contains means and standard errors of the key variables, namely the two forms of pessimism, proxies of short-term integration and savings, as well as important socio-demographic characteristics (age, years of education, gender, being married, number of children and cohort of arrival). Arrival cohorts are defined at the decade level: 1982-90; 1991-2000 and 2001-2010. Integration efforts i are proxied by the country of origin of the newspapers read by migrants on a scale from 1 (only from the country of origin) to 5 (only from Germany). The saving rate is the proportion of savings in the household's monthly income.

[28] Unfortunately, the question about predicted pay raise in GSOEP is asked at a different time horizon (2 years) than the question about predicted life satisfaction (5 years).

[29] The GSOEP includes information on several drop-out studies which were designed to understand reasons for non-response: attrition, mobility (including emigration), death etc. In addition, in case the respondent(s) can not be found at their known address, the pollster may inquire with neighbours about their possible whereabouts. For additional details on the identification of emigrants, see Kroh and Kröger (2020).

Table 1: Descriptive statistics

	(1)			(2)			(3)			(4)	
	All			UnexpS			non_unexpS			Difference	
	mean	sd		mean	sd		mean	sd		(2)-(3)	t
Temp. intentions	0.49	(0.50)	181	1.00	(0.00)	61	0.23	(0.42)	120		
Still in Germany	0.77	(0.42)	181	1.00	(0.00)	61	0.66	(0.48)	120		
UnexpS (years)[1]	10.81	(8.19)	53	15.23	(6.16)	35	2.22	(3.37)	18	13.01***	(9.94)
LS-pessimism	-0.99	(1.87)	145	-0.68	(1.50)	50	-1.16	(2.02)	95	0.48	(1.61)
W-pessimism	-0.17	(0.64)	82	0.03	(0.62)	35	-0.32	(0.63)	47	0.35*	(2.50)
Age	29.22	(10.29)	181	29.18	(11.00)	61	29.24	(9.96)	120	-0.06	(-0.04)
Female	0.54	(0.50)	181	0.48	(0.50)	61	0.57	(0.50)	120	-0.09	(-1.16)
Married	0.81	(0.39)	181	0.80	(0.40)	61	0.82	(0.39)	120	-0.01	(-0.21)
Children	0.64	(0.99)	181	0.51	(0.79)	61	0.70	(1.08)	120	-0.19	(-1.36)
Education (years)	10.97	(2.67)	158	10.59	(2.64)	55	11.17	(2.67)	103	-0.58	(-1.32)
82-90 cohort	0.23	(0.42)	181	0.25	(0.43)	61	0.23	(0.42)	120	0.02	(0.31)
91-00 cohort	0.59	(0.49)	181	0.56	(0.50)	61	0.60	(0.49)	120	-0.04	(-0.54)
Newspapers	2.13	(1.31)	60	1.96	(1.19)	23	2.24	(1.38)	37	-0.29	(-0.85)
Saving share	0.05	(0.09)	127	0.06	(0.09)	46	0.05	(0.10)	81	0.01	(0.56)
Observations	181			61			120			181	

Notes: [1]The number of years of unexpected staying can only be computed for immigrants whose initial intention was to stay temporarily, as this question was not asked to migrants with permanent intentions. Among migrants who report this value, those who stayed are "unexpected stayers", and those who left are "expected leavers". The average value of 2.2 years thus only applies to expected leavers, who are a subset of all migrants who are not unexpected stayers (this category also includes migrants with initial permanent intentions). Also note that the value of 15.2 for unexpected stayers is a lower bound since the unexpected staying duration is right censored: as long as these migrants remain in Germany, this number continues to increase. Arrival cohorts 1982-90 and 91-2000 provide the share of respondents who arrived within a specific decade. The remaining 18% of respondents arrived between 2001 and 2010. ***,**,* denote significance at the 1, 5 and 10% level, respectively.

Table 1 shows that 49% of all migrants in our sample had temporary intentions at arrival, whereas 77% are still in Germany in 2017. Out of the 181 migrants, 61 are unexpected stayers (i.e. they have stayed in Germany although they predicted that they would have left by 2017). Among the 120 migrants who are not unexpected stayers, 23% had planned to be temporary and have indeed left, 66% planned a permanent stay and are still in Germany, and the remaining 11% are unexpected leavers (i.e. they had planned to stay permanently but have left).

The average levels of pessimism -both in terms of life satisfaction and of wages- of unexpected leavers are higher than those of other migrants, with t-statistics of the difference in means of 1.61 and 2.5 respectively. The fact that these values are negative for life satisfaction suggests that migrants are on average optimistic (which is sensible given their decision to migrate to Germany), but that unexpected stayers are less optimistic than other migrants. Unexpected stayers integrated less and saved more than other migrants in the short term, though these differences are not statistically significant. Apart from pessimism, unexpected stayers do not appear to bear significant differences in their main socio-demographic characteristics compared to other migrants.

Unexpected staying and pessimism The equation that we estimate using OLS takes the form:

$$Y_i = c + \beta Pessimism_i + \sum_j \gamma_j X_{ij} + e_i,$$

where Y_i is a binary variable which equals 1 for migrant i if they are unexpected stayers in the long term and 0 otherwise, $Pessimism_i$ is measured in the first two years since arrival in Germany and pertains to either life satisfaction ($LS\_pessimism$) or wages ($wage\_pessimism$), and X_{ij} is a set of control variables capturing migrant characteristics. Note that due to data availability constraints, we estimate the effect of each of the two pessimism measures independently.[30] The main parameter of interest β is expected to be positive as the model predicts that pessimism at arrival about future utility and wages leads to an underestimation of the probability of staying in the host country in the long term.

Tables 2 and 3 provide estimates of the determinants of unexpected staying. Column (1) controls for a set of individual characteristics (age, being married, female, number of children in the household and number of years of education). Column (2) includes in addition cohort fixed effects (defined at the decade level: 1982-1990; 1991-2000 and 2001-2010). Life events have been shown to affect individuals' duration of stay (de Groot et al. (2011); Bettin et al. (2018)). Columns (3)-(5) add controls for individual level shocks that could have affected unexpected staying. Column (3) adds as a control the difference between initial marital status and the marital status observed in the last

[30] Our sample of migrants for whom we observe both life satisfaction and wage pessimism only contains 41 observations. On this sample, we find that the two measures of pessimism have a correlation of 0.01, suggesting that when migrants have misperceptions about the destination country, they are not necessarily pessimistic about both wages and utility.

Table 2: Unexpected staying and wage pessimism

	(1)		(2)		(3)		(4)		(5)	
	UnexpS		UnexpS		UnexpS		UnexpS		UnexpS	
	b	se	b	se	b	se	b	se	b	se
W-pessim.	0.188**	(0.09)	0.203**	(0.09)	0.202**	(0.09)	0.194**	(0.09)	0.195**	(0.09)
Age	-0.004	(0.01)	-0.008	(0.01)	-0.006	(0.01)	-0.004	(0.01)	-0.003	(0.01)
Female	-0.059	(0.11)	-0.050	(0.11)	-0.046	(0.11)	-0.041	(0.11)	-0.039	(0.11)
Married	0.051	(0.12)	0.112	(0.12)	0.165	(0.14)	0.069	(0.13)	0.110	(0.15)
Children	-0.084	(0.06)	-0.053	(0.07)	-0.058	(0.07)	-0.031	(0.07)	-0.037	(0.07)
Education	-0.011	(0.02)	-0.018	(0.02)	-0.021	(0.02)	-0.020	(0.02)	-0.021	(0.02)
82-90 cohort			-0.220	(0.17)	-0.205	(0.17)	-0.233	(0.17)	-0.222	(0.18)
91-00 cohort			-0.343**	(0.16)	-0.297*	(0.17)	-0.338**	(0.16)	-0.308*	(0.17)
Δ married					0.123	(0.16)			0.083	(0.16)
Δ children							0.073	(0.06)	0.064	(0.07)
Constant	0.705**	(0.27)	1.089***	(0.33)	0.974***	(0.36)	0.931**	(0.36)	0.873**	(0.38)
Observations	89		89		89		89		89	

Notes: Unexpected stayers (with UnexpS=1) are migrants who report an initial intention to leave Germany but who stayed beyond their intended duration and who are still in Germany in 2017. ***,**,* denote significance at the 1, 5 and 10% level, respectively.

survey year. A migrant who was married at arrival (value of married=1) but single, divorced or widowed in the last survey year (value of married =0) would thus have a change in marital status= -1. Column (4) adds the change in number of children ("change nbr. children") in the migrant's household as an additional control. It is calculated as the difference between the maximum number of children ever observed in the household and the initial number of children. For a respondent who is living with two children in the last survey period, but who lived with three children five years earlier and who entered GSOEP without children, the value of "change nbr. children" would thus be 3 (as we have 3-0). Column (5) jointly adds the controls introduced separately in columns (3) and (4). In both Table 2 and Table 3, the coefficient of pessimism is positive and significant, which highlights a positive association pessimism (about both life satisfaction and wages) and unexpected staying. Results are very robust to the different sets of controls.

Appendix B.1 provides results of linear regressions on the link between wage pessimism and pessimism about life satisfaction on (1) stating temporary intentions at arrival and (2) still being in Germany in 2017. Migrants who were pessimistic about their wages did not form different intentions ex ante, but tend to stay more in Germany in the long term. As a result, pessimism about wages is positively associated with unexpected staying, as highlighted in Table 2. Migrants who were pessimistic about their life satisfaction tend to formulate temporary intentions ex ante, but do not appear to have different actual return behaviors than non-pessimistic migrants. Since their intentions are more often temporary, these migrants are more likely to be unexpected stayers (see

Table 3: Unexpected staying and pessimism about life satisfaction

	(1)		(2)		(3)		(4)		(5)	
	UnexpS		UnexpS		UnexpS		UnexpS		UnexpS	
	b	se	b	se	b	se	b	se	b	se
LS-pessim.	0.046**	(0.02)	0.041*	(0.02)	0.039*	(0.02)	0.038*	(0.02)	0.037	(0.02)
Age	0.001	(0.00)	0.001	(0.00)	0.002	(0.00)	0.003	(0.00)	0.003	(0.00)
Female	-0.117	(0.08)	-0.089	(0.08)	-0.082	(0.08)	-0.082	(0.08)	-0.077	(0.08)
Education	-0.016	(0.02)	-0.012	(0.02)	-0.011	(0.02)	-0.010	(0.02)	-0.009	(0.02)
Married	-0.063	(0.11)	-0.089	(0.11)	-0.007	(0.13)	-0.100	(0.11)	-0.030	(0.13)
Children	-0.082	(0.05)	-0.113**	(0.05)	-0.107**	(0.05)	-0.085	(0.06)	-0.085	(0.06)
82-90 cohort			0.357**	(0.17)	0.352**	(0.17)	0.342**	(0.17)	0.341**	(0.17)
91-00 cohort			-0.017	(0.10)	-0.000	(0.10)	-0.010	(0.10)	0.003	(0.10)
Δ married					0.142	(0.12)			0.119	(0.12)
Δ children							0.047	(0.04)	0.040	(0.04)
Constant	0.716***	(0.26)	0.649**	(0.28)	0.533*	(0.30)	0.496	(0.31)	0.423	(0.32)
Observations	141		141		141		141		141	

Notes: Unexpected stayers (with UnexpS=1) are migrants who report an initial intention to leave Germany but who stayed beyond their intended duration and who are still in Germany in 2017. ***,**,* denote significance at the 1, 5 and 10% level, respectively.

Table 3). Appendix B shows how short-term decisions about integration and savings are affected by pessimism and temporary intentions at arrival, as well as by other covariates.

In Appendix B.2, we provide regressions of short-term decisions -integration and savings- on pessimism, controlling for migrant characteristics and initial return intentions. Integration, proxied by the type of journal that migrants read, is negatively associated with both measures of pessimism, though these estimates are not significantly different from 0. We do not find any link between pessimism and migrants' monthly saving share.

Summing up, our empirical investigation provides support for our theoretical results, as pessimism about utility and future wages is positively and significantly associated with unexpected staying. A more thorough empirical analysis of the mechanisms would require richer data than we currently have. Direct tests of mispredicted country-specific utilities would require country-specific (contingent) life satisfaction predictions. Also, the analysis would benefit from larger sample sizes and more refined measures of pessimism about labor market prospects. Finally, our measure of pessimism is one of greater pessimism about the destination than the origin country – thus migrants who are optimistic about the destination country may still be comparatively pessimistic *if* they are even more optimistic about their origin country. Thus, ideal data would cover not only beliefs about the destination country, but also about the country of origin.

8 Conclusion

Using a simple framework, we identify pessimistic misperceptions by migrants about the destination country at arrival as potential causes for unexpected staying. We explore biases that systematically give rise to such pessimism, such as projection bias and narrow bracketing, and others that do not, such as present bias. Empirically, we find a positive association between pessimism at arrival in the destination country and unexpected staying in the long term. Specifically, ex ante pessimism about life satisfaction and future wages (compared to their actual realization a few years later) leads to more unexpected staying on average, as well as lower integration and higher savings.

These findings have potentially important policy implications. While populations and governments of host countries often fear that migrants are naively optimistic and eventually do not integrate in the labor market and in society, unexpected staying and low integration may instead result from relative pessimism about the host country at arrival. This suggests that, conditional on having migrated, migrants and the host society at large would benefit from clearer signals about the costs and benefits of integration and about long term prospects. Such improvements could be achieved by more active information campaigns about the benefits of integration and actual perspectives of long-term stays, as well as integration policies such as early cohesion-enhancing community associations.

This study calls for future research in at least two dimensions. First, in our empirical results, identification could be improved if one could exploit plausibly exogenous changes in pessimism. Such exogenous variation may be model-specific: for projection bias, the main assumption is that integration and utility of life in the destination country cause mispredictions. Thus policies or field experiments that exogenously change either of these by helping with integration early on would serve to identify it.[31] Of course, such designs would face the challenging constraint that unexpected staying requires a long-term data collection process. More generally, better data that directly elicits the beliefs and predictions of migrants would help to measure misperceptions, and provide better evidence for the actual degree of unexpected staying.

Second, while the causes of unexpected staying that we identify emerge from general misperceptions about any outcome that affects the return decision, we focused on behavioral biases as the source of these misperceptions. The primary reason is that any misperception requires some level of misinference or misprediction, which in turn makes biases likely candidates. Nonetheless, there are certainly situations where inference is sufficiently hard so that institutions and networks become the primary source of information. If the institutions or networks migrants rely on are biased, and migrants have little outside information, their predictions are likely to be inaccurate.

[31] Another approach would be to follow Odermatt and Stutzer (2019), who use events that have temporary but large effects on life satisfaction to identify projection bias. We have not been able to identify such shocks that would affect the level of integration of migrants.

Enriching our framework to explicitly explore how institutions and networks can lead to pessimistic misperceptions could be an interesting extension.

References

Achenbach, R. (2017). Return Migration Decision Making: Theoretical Considerations. In *Return Migration Decisions*, pages 27–77. Springer Fachmedien Wiesbaden.

Adda, J., Dustmann, C., and GÃűrlach, J.-S. (2019). The Dynamics of Return Migration, Human Capital Accumulation, and Wage Assimilation. unpublished.

Adda, J., Dustmann, C., and Mestres, J. (2006). A Dynamic Model of Return Migration. *Essays on Temporary Migration*.

Agyeman, E. (2011). Holding on to European residence rights vs the desire to return to origin country: a study of the return intentions and return constraints of Ghanaian migrants in Vic. *Migraciones*, 30:135–159.

Akwasi Agyeman, E. and Fernández Garcia, M. (2016). Connecting Return Intentions and Home Investment: the Case of Ghanaian Migrants in Southern Europe. *Journal of International Migration and Integration*, 17(3):745–759.

Alberts, H. C. and Hazen, H. D. (2005). "There are always two voices...": International students' intentions to stay in the United States or return to their home countries.

Amit, K. (2010). Determinants of life satisfaction among immigrants from Western countries and from the FSU in Israel. *Social Indicators Research*, 96(3):515–534.

Amit, K. and Bar-Lev, S. (2014). Immigrants' Sense of Belonging to the Host Country: The Role of Life Satisfaction, Language Proficiency, and Religious Motives. *Social Indicators Research*, 124(3):947–961.

Ammassari, S. (2004). From nation-building to entrepreneurship: The impact of élite return migrants in Côte d'Ivoire and Ghana. *Population, Space and Place*, 10(2):133–154.

Angelini, V., Casi, L., and Corazzini, L. (2015). Life satisfaction of immigrants: does cultural assimilation matter? *Journal of Population Economics*, 28(3):817–844.

Ariely, D. and Loewenstein, G. (2006). The heat of the moment: The effect of sexual arousal on sexual decision making. *Journal of Behavioral Decision Making*, 19(2):87–98.

Augenblick, N. and Rabin, M. (2018). An Experiment on Time Preference and Misprediction in Unpleasant Tasks. *Review of Economic Studies*.

Badger, G. J., Bickel, W. K., Giordano, L. A., Jacobs, E. A., Loewenstein, G., and Marsch, L. (2007). Altered states: The impact of immediate craving on the valuation of current and future opioids. *Journal of Health Economics*, 26(5):865–876.

Baseler, T. (2020). Hidden Income and the Perceived Returns to Migration: Experimental Evidence from Kenya. *mimeo*.

Bettin, G., Cela, E., and Fokkema, T. (2018). Return intentions over the life course: Evidence on the effects of life events from a longitudinal sample of first- and second-generation Turkish migrants in Germany. *Demographic Research*, 39(1):1009–1038.

Borjas, G. J. and Bratsberg, B. (1996). Who leaves? The outmigration of the foreign-born. *Review of Economics and Statistics*.

Bossavie, L., Görlach, J.-S., Özden, C., and Wang, H. (2020). Temporary Migration and Entrepreneurship in Bangladesh. *mimeo*.

Bushong, B. and Gagnon-Bartsch, T. (2020). An experiment in interpersonal projection bias. Working Paper.

Chabé-Ferret, B., Machado, J., and Wahba, J. (2018). Remigration intentions and migrants' behavior. *Regional Science and Urban Economics*, 68:56–72.

Chetty, R. (2015). Behavioral economics and public policy: A pragmatic perspective. *American Economic Review*, 105(5):1–33.

Chiswick, B. and Miller, P. (2012). Negative and Positive Assimilation, Skill Transferability, and Linguistic Distance. *Journal of Human Capital*, 6(1):35–55.

Coulter, R. (2013). Wishful Thinking and the Abandonment of Moving Desires over the Life Course. *Environment and Planning A: Economy and Space*, 45(8):1944–1962.

de Groot, C., Mulder, C. H., Das, M., and Manting, D. (2011). Life Events and the Gap between Intention to Move and Actual Mobility. *Environment and Planning A: Economy and Space*, 43(1):48–66.

De Weerdt, J., Genicot, G., and Mesnard, A. (2019). Asymmetry of Information within Family Networks. *Journal of Human Resources*, 54(1):225–254.

Delpierre, M. and Verheyden, B. (2014). Remittances, savings and return migration under uncertainty. *IZA Journal of Migration*, 3(1):1–43.

Dohmen, T. J., Falk, A., Huffman, D., Marklein, F., and Sunde, U. (2009). Biased probability judgment: Evidence of incidence and relationship to economic outcomes from a representative sample. *Journal of Economic Behavior & Organization*, 72(3):903–915.

Dustmann, C. (1999). Temporary migration, human capital, and language fluency of migrants. *Scandinavian Journal of Economics*, 101(2):297–314.

Dustmann, C. and Fabbri, F. (2003). Language proficiency and labour market performance of immigrants in the UK. *Economic Journal*, 113(489):695–717.

Dustmann, C. and Görlach, J.-S. (2016). The Economics of Temporary Migrations. *Journal of Economic Literature*, 54(1):98–136.

Dustmann, C. and Mestres, J. (2010a). Remittances and temporary migration. *Journal of Development Economics*, 92(1):62–70.

Dustmann, C. and Mestres, J. (2010b). Savings, asset holdings, and temporary migration. *Annals of Economics and Statistics/Annales d Économie et de Statistique*, pages 289–306.

Dustmann, C. and Soest, A. V. (2002). Language and the Earnings of Immigrants. *ILR Review*, 55(3):473–492.

Eliaz, K. and Spiegler, R. (2015). Beyond ellison's matrix: New directions in behavioral industrial organization. *Review of Industrial Organization*, 47(3):259–272.

Ellis, A. and Freeman, D. J. (2020). Revealing choice bracketing. *arXiv preprint arXiv:2006.14869*.

Enke, B. (2020). What You See Is All There Is*. *The Quarterly Journal of Economics*, 135(3):1363–1398.

Esponda, I. (2008). Behavioral equilibrium in economies with adverse selection. *American Economic Review*, 98(4):1269–91.

Exley, C. L. and Kessler, J. B. (2018). Equity concerns are narrowly framed. Technical report, National Bureau of Economic Research.

Eyster, E. and Rabin, M. (2005). Cursed equilibrium. *Econometrica*, 73(5):1623–1672.

Fallucchi, F. and Kaufmann, M. (2020). Narrow bracketing in work choices. Working Paper.

Geurts, N. and Lubbers, M. (2017). Dynamics in intention to stay and changes in language proficiency of recent migrants in the Netherlands. *Journal of Ethnic and Migration Studies*, 43(7):1045–1060.

Gould, W., McManus, W., and Welch, F. (1983). Earnings of Hispanic Men: The Role of English Language Proficiency . *Journal of Labor Economics*, 1(2):101–30.

Guven, C. and Islam, A. (2015). Age at migration, language proficiency, and socioeconomic outcomes: evidence from Australia. *Demography*, 52(2):513–542.

Handel, B. and Schwartzstein, J. (2018). Frictions or mental gaps: what's behind the information we (don't) use and when do we care? *Journal of Economic Perspectives*, 32(1):155–78.

Hoffman, M. and Burks, S. V. (2020). Worker overconfidence: Field evidence and implications for employee turnover and firm profits. *Quantitative Economics*, 11(1):315–348.

Hooijen, I., Meng, C., and Reinold, J. (2020). Be prepared for the unexpected: The gap between (im)mobility intentions and subsequent behaviour of recent higher education graduates. *Population, Space and Place*, 26(5).

Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):263–292.

Kirdar, M. (2013). Source country characteristics and immigrants optimal migration duration decision. *IZA Journal of Migration*, 2(1):8.

Kley, S. (2017). Facilitators and constraints at each stage of the migration decision process. *Population Studies*, 71(sup1):35–49.

Kley, S. A. and Mulder, C. H. (2010). Considering, planning, and realizing migration in early adulthood. The influence of life-course events and perceived opportunities on leaving the city in Germany. *Journal of Housing and the Built Environment*, 25(1):73–94.

Kogan, I., Shen, J., and Siegert, M. (2018). What Makes a Satisfied Immigrant? Host-Country Characteristics and Immigrants' Life Satisfaction in Eighteen European Countries. *Journal of Happiness Studies*, 19(6):1783–1809.

Kroh, M. and Kröger, H. (2020). SOEP-Core v35 LIFESPELL: Information on the Pre-and Post-Survey History of SOEP-Respondents. *SOEP Survey Papers: Series D*, 887.

Krueger, A. B. and Mueller, A. I. (2016). A contribution to the empirics of reservation wages. *American Economic Journal: Economic Policy*, 8(1):142–179.

Lochmann, A., Rapoport, H., and Speciale, B. (2019). The effect of language training on immigrantsâĂŹ economic integration: Empirical evidence from France. *European Economic Review*, 113(C):265–296.

Loewenstein, G., O'Donoghue, T., and Rabin, M. (2003). Projection bias in predicting future utility. *Quarterly Journal of Economics*, 118(4):1209–1248.

Lu, M. (1999). Do people move when they say they will? Inconsistencies in individual migration behavior. *Population and Environment*, 20(5):467–488.

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

Mullainathan, S., Schwartzstein, J., and Congdon, W. J. (2012). A reduced-form approach to behavioral public finance. *Annu. Rev. Econ.*, 4(1):511–540.

Odermatt, R. and Stutzer, A. (2019). Mis-predicted subjective well-being following life events. *Journal of the European Economic Association*, 17(1):245–283.

O'Donoghue, T. and Rabin, M. (1999). Doing It Now or Later. *The American Economic Review*, 89(1):103–124.

Pinger, P. R. et al. (2017). Predicting experimental choice behavior and life outcomes from a survey measure of present bias. *Economics Bulletin*, 37(3):2162–2172.

Rabin, M. and Weizsäcker, G. (2009). Narrow bracketing and dominated choices. *American Economic Review*, 99(4):1508–1543.

Read, D., Loewenstein, G., Rabin, M., Keren, G., and Laibson, D. (1999). Choice bracketing. In *Elicitation of preferences*, pages 171–202. Springer.

Redelmeier, D. A. and Tversky, A. (1992). On the framing of multiple prospects. *Psychological Science*, 3(3):191–193.

Schoorl, J. (2011). Arbeidsmigratie door de jaren heen. *Demos*, 27(10).

Seshan, G. and Zubrickas, R. (2017). Asymmetric Information about Migrant Earnings and Remittance Flows. *World Bank Economic Review*, 31(1):24–43.

Shrestha, M. (2020). Get rich or die trying: Perceived earnings, perceived mortality rate and the value of a statistical life of potential work-migrants from Nepal. *mimeo*.

Sinning, M. (2011). Determinants of savings and remittances: empirical evidence from immigrants to Germany. *Review of Economics of the Household*, 9(1):45–67.

Spinnewijn, J. (2015). Unemployed but optimistic: optimal insurance design with biased beliefs. *Journal of the European Economic Association*, 13(1):130–167.

Steiner, V. and Velling, J. (1994). Re-migration behavior and expected duration of stay of guest workers in germany. In *The economic consequences of immigration to Germany*, pages 101–119. Springer.

van Baalen, B. and Muller, T. (2008). Return intentions of temporary migrants: the case of Germany. *CEPR proceedings of the second conference on transnationality of migrants.*

van Dalen, H. P. and Henkens, K. (2013). Explaining emigration intentions and behaviour in the Netherlands, 2005-10. *Population Studies*, 67(2):225–241.

van den Berg, G. J. and Weynandt, M. A. (2013). Explaining Differences Between the Expected and Actual Duration Until Return Migration: Economic Changes. *Schmollers Jahrbuch*, 133(2):249–261.

Van Tubergen, F. and Kalmijn, M. (2009). Language proficiency and usage among immigrants in the Netherlands: Incentives or opportunities? *European Sociological Review*, 25(2):169–182.

Waldorf, B. (1995). Determinants of international return migration intentions. *Professional Geographer*, 47(2):125–136.

Ward, Z. (2017). Birds of passage: Return migration, self-selection and immigration quotas. *Explorations in Economic History*, 64:37–52.

Weynandt, M. A. (2014). *Essays in Applied Econometrics*. PhD thesis, University of Mannheim.

A Appendix: Proofs

A.1 Two lemmas introducing Proposition 1

First, let us introduce two useful lemmas which describe the structure of migrant's programme.

Lemma 1. *Let $u(c,i)$ be separable in c and i. Then for any given (s,i), additional integration i reduces the reservation wage $w_R(s,i)$, and additional savings s increase the reservation wage:*

$$\begin{aligned}\frac{\partial w_R(s,i)}{\partial i} &= -\frac{\partial v^d(s,i;w_R)/\partial i}{\partial v^d(s,i;w_R)/\partial s} < 0,\\ \frac{\partial w_R(s,i)}{\partial s} &= \frac{\partial v^o(s)/\partial s}{\partial v^d(s,i;w_R)/\partial s} - 1 > 0.\end{aligned}$$

In the space (i,s), the locus of points (s,i) such that $w_R(s,i) = \overline{w}$ is (i) increasing for all constant $\overline{w}$ and (ii) concave.

Proof. Note that $w_R(s,i)$ is defined by the implicit function $v^d(s,i;w_R) - v^o(s) = 0$, or equivalently $u^d\left(c^d(s), I^d(i)\right) - u^o(c^o(s), I^o) = 0$, where $c^d = w_R + s$, $I^d = I_1 + i$, $c^o = (w^o + s)x$, and I^o is exogenous. Applying the implicit function theorem to this equality, one obtains the two equalities stated in the Lemma:

$$\begin{aligned}\frac{\partial w_R(s,i)}{\partial i} &= -\frac{\frac{\partial u^d\left(c^d,I^d;w_R\right)}{\partial I^d}}{\frac{\partial u^d\left(c^d,I^d;w_R\right)}{\partial c^d}} = -\frac{\frac{\partial v^d(s,i;w_R)}{\partial i}}{\frac{\partial v^d(s,i;w_R)}{\partial s}},\\ \frac{\partial w_R(s,i)}{\partial s} &= -\frac{\frac{\partial u^d\left(c^d,I^d;w_R\right)}{\partial c^d} - x\frac{\partial u^o(c^o,I^o)}{\partial c^o}}{\frac{\partial u^d\left(c^d,I^d;w_R\right)}{\partial c^d}} = \frac{\frac{\partial v^o(s;w_R)}{\partial s} - \frac{\partial v^d(s,i;w_R)}{\partial s}}{\frac{\partial v^d(s,i;w_R)}{\partial s}}\end{aligned}$$

Note that period-2 utility is increasing in i and in s, which implies that $\frac{\partial w_R(s,i)}{\partial i} < 0$. To show that $\frac{\partial w_R(s,i)}{\partial s} > 0$, we prove that $\partial v^o(s)/\partial s > \partial v^d(s,i;w_R)/\partial s$, or equivalently that $x\partial u^o(c^o,I^o)/\partial c^o > \partial u^d\left(c^d,I^d;w_R\right)/\partial c^d$. First note that $x > 1$by assumption. Second, by definition of w_R, these derivatives are compared for consumption and integration levels such that $u^d\left(c^d,I^d\right) = u^o(c^o,I^o)$. Since for all i, $I^o \geq I^d$, it must be that $c^d \geq c^o$ (otherwise the migrant could not be indifferent between the two locations). Therefore, if $u(c,I)$ is separable, $c^d \geq c^o$ implies that $\partial u^o(c^o,I^o)/\partial c^o > \partial u^d\left(c^d,I^d;w_R\right)/\partial c^d$. Note that if $u(c,I)$ has positive cross partial derivatives, then the higher I^o reinforces the result.

Also note that in the space (i,s), the locus of points which yield the same $w_R(s,i)$ is increasing and concave:

$$\begin{aligned}\frac{\partial s}{\partial i}|_{w_R cst} &= -\frac{\frac{\partial w_R(s,i)}{\partial i}}{\frac{\partial w_R(s,i)}{\partial s}} = -\frac{\frac{\partial v^d(s,i;w_R)}{\partial i}}{\frac{\partial v^d(s,i;w_R)}{\partial s} - \frac{\partial v^o(s;w_R)}{\partial s}} = \frac{\frac{\partial v^d(s,i;w_R)}{\partial i}}{\frac{\partial v^o(s;w_R)}{\partial s} - \frac{\partial v^d(s,i;w_R)}{\partial s}} > 0,\\ \frac{\partial^2 s}{\partial i^2}|_{w_R cst} &= \frac{\frac{\partial^2 v^d(s,i;w_R)}{\partial i^2}}{\frac{\partial v^o(s,i;w_R)}{\partial s} - \frac{\partial v^d(s,i;w_R)}{\partial s}} < 0.\end{aligned}$$

□

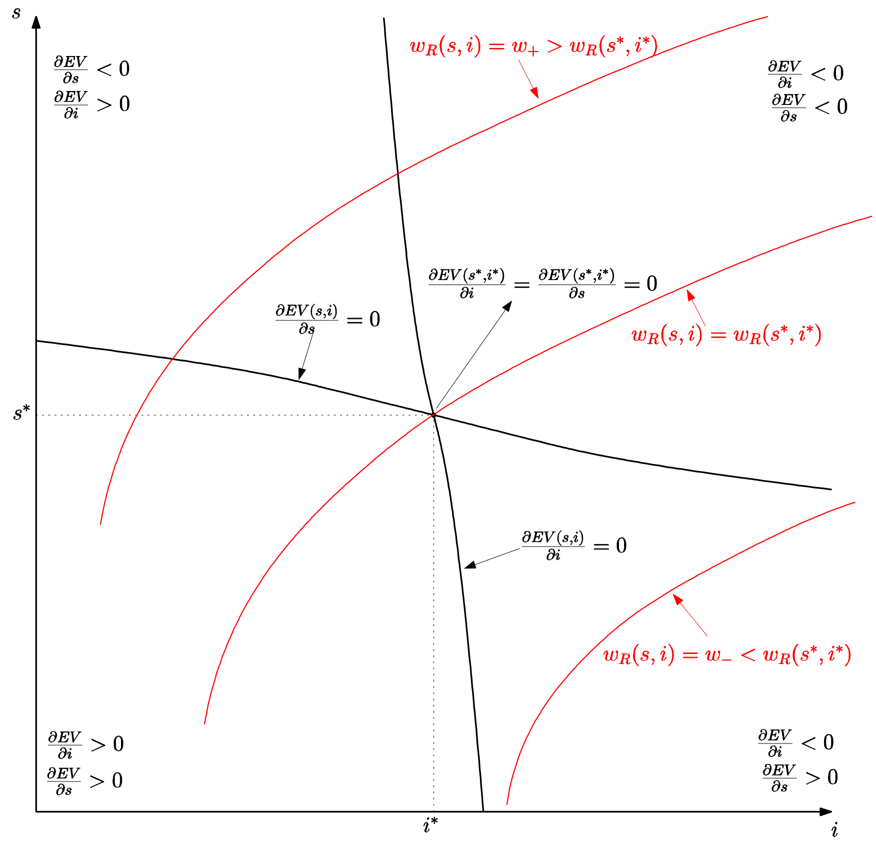

Figure 1: Representation of the reservation wage and first order conditions in the (i,s) space

The fact that the iso-w_R curve is increasing and concave is illustrated by the red curves in Figure 1. The red curve at the top of Figure 1 represents all the combinations of (s,i) such that $w_R(s,i) = w_+$, where w_+ is some positive constant. The red curve at the bottom of Figure 1 represents all the combinations of (s,i) such that $w_R(s,i) = w_-$, where (by $\frac{\partial w_R(s,i)}{\partial i} < 0$ and $\frac{\partial w_R(s,i)}{\partial s} > 0$) $w_- < w_+$. The central red curve represents all the combinations of (s,i) which generate the same reservation wage as $w_R(s^*,i^*)$, the reservation wage obtained from optimal actions (s^*,i^*). Such optimal actions are defined by Lemma 2.

Lemma 2. *Let* $(s^*,i^*) = \arg\max_{s,i} EV(s,i)$, *where* $EV(s,i) = EV\left(s,i;v_1,v^d,v^o,F\right)$, *and* $\left(\widetilde{s},\widetilde{i}\right) = \arg\max_{s,i} \widetilde{EV}(s,i)$, *where* $\widetilde{EV}(s,i) = EV\left(s,i;\widetilde{v}_1,\widetilde{v}^d,\widetilde{v}^o,\widetilde{F}\right)$.[32] *Then* (s^*,i^*) *must satisfy*

$$\frac{\partial EV}{\partial i} = \frac{\partial v_1(s,i)}{\partial i} + \delta \int_{w_R}^{\infty} \frac{\partial v^d(s,i)}{\partial i} f\left(w^d\right) dw^d = 0,$$

$$\frac{\partial EV}{\partial s} = \frac{\partial v_1(s,i)}{\partial s} + \delta \int_{0}^{w_R} \frac{\partial v^o(s)}{\partial s} f\left(w^d\right) dw^d + \delta \int_{w_R}^{\infty} \frac{\partial v^d(s,i)}{\partial s} f\left(w^d\right) dw^d = 0,$$

[32] See 3.2 for the full expression of EV.

and $\frac{\partial^2 EV}{\partial i^2} < 0$, $\frac{\partial^2 EV}{\partial s^2} < 0$ *and* $\frac{\partial^2 EV}{\partial s \partial i} < 0$, *where under separable preferences* $u(c,i)$,

$$\begin{aligned}
\frac{\partial^2 EV}{\partial i^2} &= \frac{\partial^2 v_1(s,i)}{\partial i^2} + \delta \int_{w_R}^{\infty} \frac{\partial^2 v^d(s,i)}{\partial i^2} f\left(w^d\right) dw^d + 2\delta \frac{\left(\frac{\partial v^d(s,i;w_R)}{\partial i}\right)^2}{\frac{\partial v^d(s,i;w_R)}{\partial s}} f(w_R), \\
\frac{\partial^2 EV}{\partial s^2} &= \frac{\partial^2 v_1(s,i)}{\partial s^2} + \delta \left(\int_0^{w_R} \frac{\partial^2 v^o(s)}{\partial s^2} f\left(w^d\right) dw^d + \int_{w_R}^{\infty} \frac{\partial^2 v^d(s,i)}{\partial s^2} f\left(w^d\right) dw^d \right) \\
&\quad + 2\delta \frac{\left(\frac{\partial v^o(s)}{\partial s} - \frac{\partial v^d(s,i)}{\partial s}\right)^2}{\frac{\partial v^d(s,i;w_R)}{\partial s}} f(w_R), \\
\frac{\partial^2 EV}{\partial s \partial i} &= -2\delta \frac{\frac{\partial v^d(s,i;w_R)}{\partial i}\left(\frac{\partial v^o(s;w_R)}{\partial s} - \frac{\partial v^d(s,i;w_R)}{\partial s}\right)}{\frac{\partial v^d(s,i;w_R)}{\partial s}} f(w_R).
\end{aligned}$$

Proof. In this proof, we first derive the first order conditions stated in Lemma 2, and then show the second and cross derivatives. Since $(s^*, i^*) = \arg\max_{s,i} EV(s,i)$, (s^*, i^*) must satisfy the following first order conditions:

$$\begin{aligned}
\frac{\partial EV}{\partial i} &= \frac{\partial v_1(s,i)}{\partial i} + \delta \int_{w_R}^{\infty} \frac{\partial v^d(s,i)}{\partial i} f\left(w^d\right) dw^d \\
&\quad + \delta \frac{\partial w_R(s,i)}{\partial i} \overbrace{\left(v^o(s;w_R) - v^d(s,i;w_R)\right)}^{=0 \text{ by definition of } w_R} f(w_R), \\
\frac{\partial EV}{\partial s} &= \frac{\partial v_1(s,i)}{\partial s} + \delta \int_0^{w_R} \frac{\partial v^o(s)}{\partial s} f\left(w^d\right) dw^d + \delta \int_{w_R}^{\infty} \frac{\partial v^d(s,i)}{\partial s} f\left(w^d\right) dw^d \\
&\quad + \delta \frac{\partial w_R(s,i)}{\partial s} \overbrace{\left(v^o(s;w_R) - v^d(s,i;w_R)\right)}^{=0 \text{ by definition of } w_R} f(w_R).
\end{aligned}$$

Note that for (s^*, i^*) to be a local maximum, the following conditions must hold:

$$\frac{\partial^2 EV}{\partial i^2} = \frac{\partial^2 v_1(s,i)}{\partial i^2} + \delta \int_{w_R}^{\infty} \frac{\partial^2 v^d(s,i)}{\partial i^2} f\left(w^d\right) dw^d + 2\delta \left(-\frac{\partial w_R(s,i)}{\partial i}\right) \frac{\partial v^d(s,i;w_R)}{\partial i} f(w_R) < 0.$$

Using Lemma 1 for $\frac{\partial w_R(s,i)}{\partial i}$, one obtains the final expression of $\frac{\partial^2 EV}{\partial i^2}$.

$$\begin{aligned}
\frac{\partial^2 EV}{\partial s^2} &= \frac{\partial^2 v_1(s,i)}{\partial s^2} + \delta \int_0^{w_R} \frac{\partial^2 v^o(s)}{\partial s^2} f\left(w^d\right) dw^d + \delta \int_{w_R}^{\infty} \frac{\partial^2 v^d(s,i)}{\partial s^2} f\left(w^d\right) dw^d \\
&\quad + 2\delta \frac{\partial w_R(s,i)}{\partial s} \left(\frac{\partial v^o(s)}{\partial s} - \frac{\partial v^d(s,i)}{\partial s}\right) f(w_R) < 0.
\end{aligned}$$

Using Lemma 1 for $\frac{\partial w_R(s,i)}{\partial s}$, one obtains the final expression of $\frac{\partial^2 EV}{\partial s^2}$.

Also, to be a maximum, (s^*, i^*) must satisfy $\frac{\partial^2 EV}{\partial i^2}\frac{\partial EV}{\partial s} - \left(\frac{\partial^2 EV}{\partial s \partial i}\right)^2 > 0$. Let us thus derive the cross derivative, which we obtain by deriving $\frac{\partial EV}{\partial s}$ with respect to i:

$$\begin{aligned}\frac{\partial^2 EV}{\partial s \partial i} &= \frac{\partial v_1(s,i)}{\partial s \partial i} + \delta \frac{\partial w_R(s,i)}{\partial i}\frac{\partial v^o(s;w_R)}{\partial s} f(w_R) \\ &+\delta \int_{w_R}^{\infty} \frac{\partial^2 v^d(s,i)}{\partial s \partial i} f\left(w^d\right) dw^d - \delta \frac{\partial w_R(s,i)}{\partial i}\frac{\partial v^d(s,i;w_R)}{\partial s} f(w_R) \\ &+\delta \frac{\partial w_R(s,i)}{\partial s}\left(-\frac{\partial v^d(s,i;w_R)}{\partial i}\right) f(w_R)\end{aligned}$$

Using separability of $v(s,i)$, this expression boils down to

$$\begin{aligned}\frac{\partial^2 EV}{\partial s \partial i} &= \delta \frac{\partial w_R(s,i)}{\partial i}\left(\frac{\partial v^o(s,i;w_R)}{\partial s} - \frac{\partial v^d(s,i;w_R)}{\partial s}\right) f(w_R) \\ &+\delta \frac{\partial w_R(s,i)}{\partial s}\left(-\frac{\partial v^d(s,i;w_R)}{\partial i}\right) f(w_R) \\ &< 0.\end{aligned}$$

Using Lemma 1, note that each of these two terms is equal to $-\delta \frac{\frac{\partial v^d(s,i;w_R)}{\partial i}\left(\frac{\partial v^o(s;w_R)}{\partial s} - \frac{\partial v^d(s,i;w_R)}{\partial s}\right)}{\frac{\partial v^d(s,i;w_R)}{\partial s}} f(w_R)$.

□

Lemma 2 can be used to illustrate the optimal choice (s^*, i^*) in Figure 1. In the space (i, s), (s^*, i^*) is characterized by the intersection of two downward-sloped curves which respectively characterize the loci $\frac{\partial EV(s,i)}{\partial i} = 0$ and $\frac{\partial EV(s,i)}{\partial s} = 0$. Note that the slope of $\frac{\partial EV(s,i)}{\partial i} = 0$ is steeper (more negative) than $\frac{\partial EV(s,i)}{\partial s} = 0$, which always holds if (s^*, i^*) is a maximum.[33] Indeed, note that the slope of $\frac{\partial EV(s,i)}{\partial i} = 0$ is $-\frac{\partial^2 EV(s,i)}{\partial i^2}/\frac{\partial^2 EV(s,i)}{\partial i \partial s}$, while the slope of $\frac{\partial EV(s,i)}{\partial s} = 0$ is $-\frac{\partial^2 EV(s,i)}{\partial s \partial i}/\frac{\partial^2 EV(s,i)}{\partial s^2}$. Thus the former is steeper than the latter at (s^*, i^*) if and only if $-\frac{\partial^2 EV(s,i)}{\partial i^2}/\frac{\partial^2 EV(s,i)}{\partial i \partial s} < -\frac{\partial^2 EV(s,i)}{\partial s \partial i}/\frac{\partial^2 EV(s,i)}{\partial s^2}$, or equivalently $\frac{\partial^2 EV}{\partial i^2}\frac{\partial^2 EV}{\partial s^2} - \left(\frac{\partial^2 EV}{\partial s \partial i}\right)^2 > 0$, which is a necessary condition for (s^*, i^*) to be a maximum.

A.2 Proof of Proposition 1

Proof. To prove that G-pessimism is sufficient, let us build on our graphical representation of the migrant's choice depicted in Figure 1, and let us introduce two new loci, $\frac{\partial E\widetilde{V}}{\partial i} = 0$ and $\frac{\partial E\widetilde{V}}{\partial s} = 0$. These two loci are represented by the blue curves in Figure 2. Let us start by explaining their position, i.e. why (1) the locus $\frac{\partial E\widetilde{V}}{\partial i} = 0$ is to the left of $\frac{\partial EV}{\partial i} = 0$ and (2) the locus $\frac{\partial E\widetilde{V}}{\partial s} = 0$ is to the right of $\frac{\partial EV}{\partial s} = 0$.

Let us take (s^*, i^*) as our reference point. Since pessimism means that $\frac{\partial E\widetilde{V}(s,i)}{\partial i} \leq \frac{\partial EV(s,i)}{\partial i}$ for all (s, i), and since (s^*, i^*) is such that $\frac{\partial EV(s^*,i^*)}{\partial i} = 0$, we conclude that $\frac{\partial E\widetilde{V}(s^*,i^*)}{\partial i} \leq 0$. Since $\frac{\partial^2 E\widetilde{V}}{\partial i^2} < 0$,

[33] Note that this is difference in slopes holds indepedently of assumptions made on the separability of the utility function.

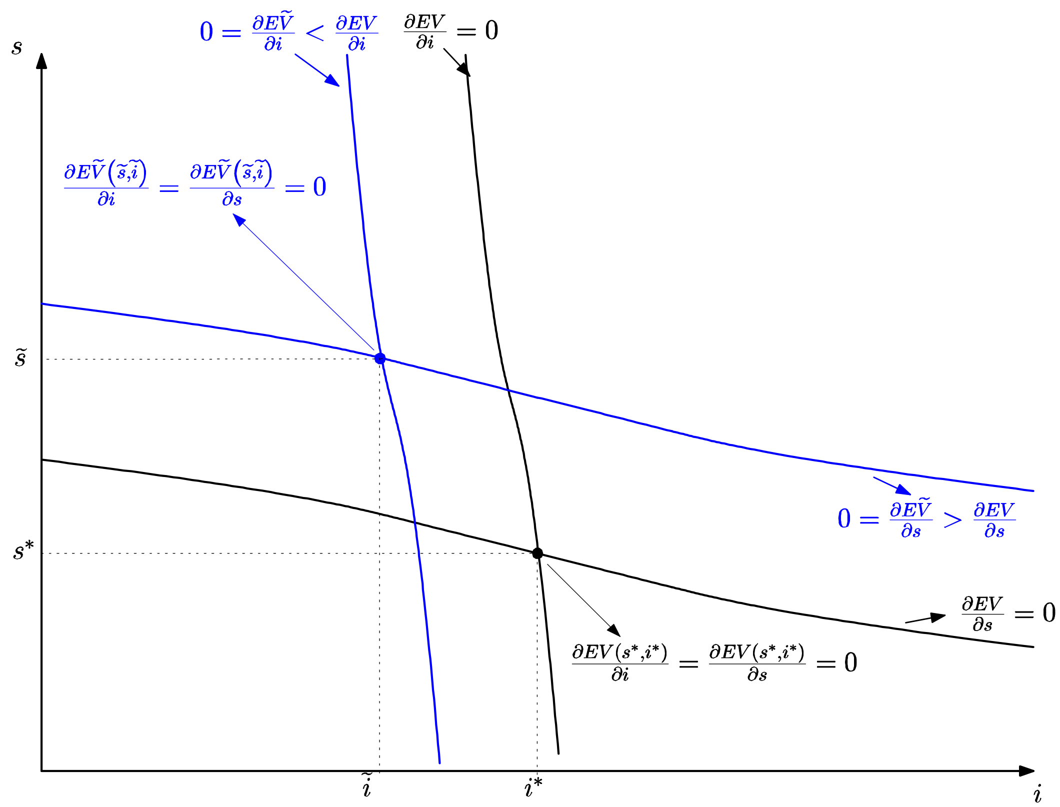

Figure 2: Representation of misperceived first order conditions and predicted choices in the (i,s) space

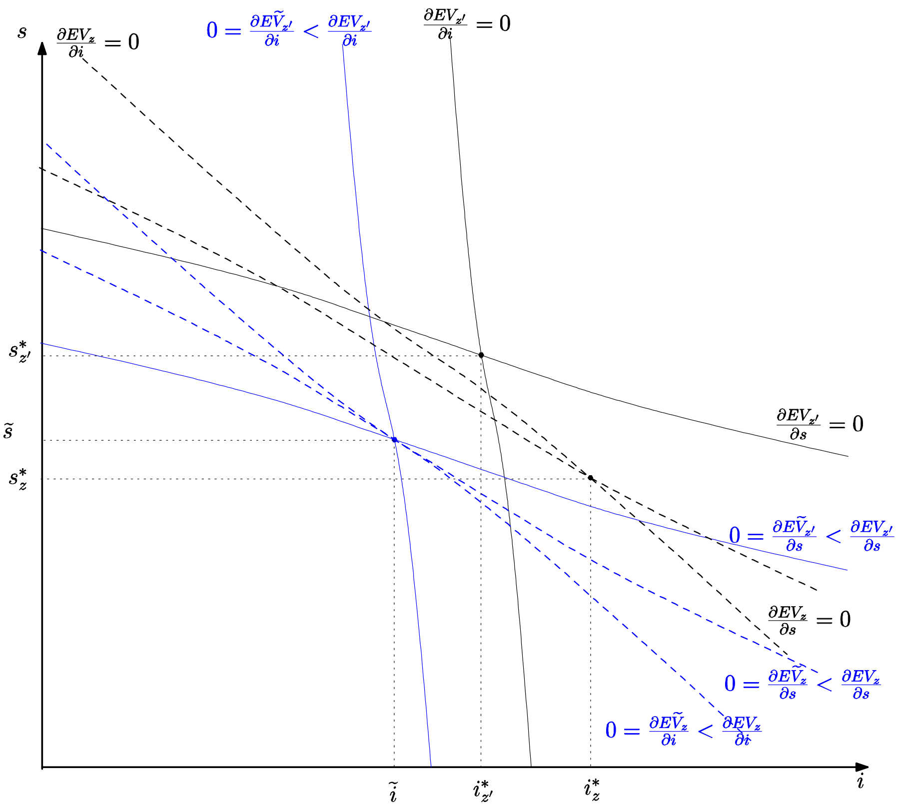

Figure 3: G-pessimism is a necessary condition for low integration for all z

for any given s, a decrease in i increases $\frac{\partial E\widetilde{V}}{\partial i}$. So the locus $\frac{\partial E\widetilde{V}(s,i)}{\partial i} = 0$ lies to the left of the locus $\frac{\partial EV(s,i)}{\partial i} = 0$. Let us apply the same reasoning with $\frac{\partial E\widetilde{V}(s,i)}{\partial s}$. Since $\frac{\partial E\widetilde{V}(s,i)}{\partial s} \geq \frac{\partial EV(s,i)}{\partial s}$ for all (s,i), and (s^*,i^*) is such that $\frac{\partial EV(s^*,i^*)}{\partial s} = 0$, we can conclude that $\frac{\partial E\widetilde{V}(s^*,i^*)}{\partial s} \geq 0$. Similarly, since $\frac{\partial^2 E\widetilde{V}}{\partial s^2} < 0$, for any given i, an increase in s decreases $\frac{\partial E\widetilde{V}}{\partial s}$. So the locus $\frac{\partial E\widetilde{V}(s,i)}{\partial s} = 0$ lies above the locus $\frac{\partial EV(s,i)}{\partial s} = 0$. Given that $\frac{\partial E\widetilde{V}}{\partial s} = 0$ lies above $\frac{\partial EV}{\partial s} = 0$ and $\frac{\partial E\widetilde{V}}{\partial i} = 0$ lies to the left of $\frac{\partial EV}{\partial s} = 0$, the intersection of the loci $\frac{\partial E\widetilde{V}}{\partial s} = 0$ and $\frac{\partial E\widetilde{V}}{\partial i} = 0$ is to the top left of (s^*, i^*). (Note that the position of the intersection of the black curves is to the upper left of that of the blue curves independently of the magnitude of their slopes, as long as they respect the basic properties described in Lemma 2 – that is, that they are downward-sloping.) Since this intersection is by definition $\left(\widetilde{s}, \widetilde{i}\right)$, we can conclude that $\widetilde{s} \geq s^*$ and $\widetilde{i} \leq i^*$.

□

A.3 Mispredicted probabilities

A.3.1 Proof of Proposition 2

Proof. The formal statement which needs to be proven is: $\tilde{F}(w^d) \geq F(w^d)$ for all $w^d \iff \widetilde{F}\left(w_R\left(\widetilde{s}, \widetilde{i}; v^d, v^o\right)\right) \geq F\left(w_R\left(\widetilde{s}, \widetilde{i}; v^d, v^o\right)\right)$ for all $\left(z, v_1, v^d, v^o, F\right)$, where $(\widetilde{s}, \widetilde{i}) = \arg\max_{s,i} EV\left(s, i; v_1, v^o, v^d, \widetilde{F}\right)$

Proving the sufficiency condition ($\Longrightarrow$) is trivial: under P-pessimism, $\tilde{F}(w^d) \geq F(w^d)$ for all w^d, thus this inequality holds for $w^d = w_R\left(\widetilde{s}, \widetilde{i}; v^d, v^o\right)$. To prove the necessity condition ($\Longleftarrow$), note that $\widetilde{F}\left(w_R\left(\widetilde{s}, \widetilde{i}; v^d, v^o\right)\right) \geq F\left(w_R\left(\widetilde{s}, \widetilde{i}; v^d, v^o\right)\right)$ must hold for any $\left(v_1, v^d, v^o, F\right)$, and thus for any possible $\widetilde{s}, \widetilde{i}$ and $w_R\left(\widetilde{s}, \widetilde{i}; v^d, v^o\right)$. Stating that this condition must hold for all possible $w_R\left(\widetilde{s}, \widetilde{i}; v^d, v^o\right)$ is formally identical to stating that $\tilde{F}(w^d) \geq F(w^d)$ for all w^d, which is our definition of P-pessimism. □

A.3.2 Proof of Proposition 3

Proof. Proposition 3 is a direct application of Proposition 1 since P-pessimism implies $\frac{\partial E\widetilde{V}}{\partial i} \leq \frac{\partial EV}{\partial i}$ and $\frac{\partial E\widetilde{V}}{\partial s} \geq \frac{\partial EV}{\partial s}$.

(1) Proof that $\frac{\partial E\widetilde{V}}{\partial i} \leq \frac{\partial EV}{\partial i}$. It suffices to show (since $\widetilde{F}$ does not directly affect v_1) in the case of mispredicted probabilities that $\frac{\partial E\widetilde{v}_2}{\partial i} - \frac{\partial Ev_2}{\partial i} \leq 0$. Using Lemma 2, this is equivalent to prove that

$$\int_{w_R}^{\infty} \frac{\partial v^d(s,i)}{\partial i}\left(\widetilde{f}\left(w^d\right) - f\left(w^d\right)\right) dw^d \leq 0.$$

Integrating this expression by parts, one obtains $\int_{w_R}^{\infty} \frac{\partial v^d(s,i)}{\partial i}\left(\widetilde{f}\left(w^d\right) - f\left(w^d\right)\right) dw^d$

$$\begin{aligned}
&= \left[\frac{\partial v^d(s,i)}{\partial i}\left(\widetilde{F}\left(w^d\right) - F\left(w^d\right)\right)\right]_{w_R}^{\infty} - \int_{w_R}^{\infty} \frac{\partial^2 v^d(s,i)}{\partial i \partial w^d}\left(\widetilde{F}\left(w^d\right) - F\left(w^d\right)\right) dw^d, \\
&= -\frac{\partial v^d(s,i;w_R)}{\partial i}\left(\widetilde{F}(w_R) - F(w_R)\right) - \int_{w_R}^{\infty} \frac{\partial^2 v^d(s,i)}{\partial i \partial w^d}\left(\widetilde{F}\left(w^d\right) - F\left(w^d\right)\right) dw^d.
\end{aligned}$$

The first term of this expression is negative by $\frac{\partial v^d}{\partial i} > 0$ and FOSD. The second term is also negative by FOSD and as long the cross partial derivative of $u(c,I)$ is non-negative.

(2) Proof that $\frac{\partial E\widetilde{V}}{\partial s} > \frac{\partial EV}{\partial s}$, which is equivalent to $\frac{\partial E\widetilde{v}_2}{\partial s} - \frac{\partial Ev_2}{\partial s} > 0$, that is:

$$\int_0^{w_R} \frac{\partial v^o}{\partial s}\left(\widetilde{f}\left(w^d\right) - f\left(w^d\right)\right) dw^d + \int_{w_R}^{\infty} \frac{\partial v^d}{\partial s}\left(\widetilde{f}\left(w^d\right) - f\left(w^d\right)\right) dw^d > 0.$$

The first term is (since v^o does not depend on w^d): $\int_0^{w_R} \frac{\partial v^o}{\partial s}\left(\widetilde{f}\left(w^d\right) - f\left(w^d\right)\right) dw^d$

$$= \frac{\partial v^o}{\partial s}\int_0^{w_R}\left(\widetilde{f}\left(w^d\right) - f\left(w^d\right)\right) dw^d = \frac{\partial v^o}{\partial s}\left(\widetilde{F}(w_R) - F(w_R)\right).$$

Let us integrate the second term by parts: $\int\limits_{w_R}^{\infty} \frac{\partial v^d}{\partial s} \left(\widetilde{f}\left(w^d\right) - f\left(w^d\right) \right) dw^d$

$$
\begin{aligned}
&= \left[\frac{\partial v^d}{\partial s} \left(\widetilde{F}\left(w^d\right) - F\left(w^d\right) \right) \right]_{w_R}^{\infty} - \int\limits_{w_R}^{\infty} \frac{\partial^2 v^d}{\partial s \partial w^d} \left(\widetilde{F}\left(w^d\right) - F\left(w^d\right) \right) dw^d \\
&= -\frac{\partial v^d\left(s, i; w_R\right)}{\partial s} \left(\widetilde{F}\left(w_R\right) - F\left(w_R\right) \right) + \int\limits_{w_R}^{\infty} \left(-\frac{\partial^2 v^d}{\partial s^2} \right) \left(\widetilde{F}\left(w^d\right) - F\left(w^d\right) \right) dw^d.
\end{aligned}
$$

Combining both terms, one obtains:

$$
\frac{\partial E\widetilde{v}_2}{\partial s} - \frac{\partial E v_2}{\partial s} = \left(\frac{\partial v^o\left(s\right)}{\partial s} - \frac{\partial v^d\left(s, i; w_R\right)}{\partial s} \right) \left(\widetilde{F}\left(w_R\right) - F\left(w_R\right) \right) + \int\limits_{w_R}^{\infty} \left(-\frac{\partial^2 v^d}{\partial s^2} \right) \left(\widetilde{F}\left(w^d\right) - F\left(w^d\right) \right) dw^d.
$$

The first term of this equation is always positive since by Lemma 1, $\left(\frac{\partial v^o(s)}{\partial s} - \frac{\partial v^d(s,i;w_R)}{\partial s} \right) > 0$. The second term is also positive by concavity of v^d and FOSD for all w^d.

(3) Extension: integration improves the distribution of wages. Let us now consider an extension of the model in which migrants can improve the distribution of wages thanks to integration: $F\left(w^d; I^d\right)$, with $\frac{\partial F\left(w^d; I^d\right)}{\partial I^d} = \frac{\partial F\left(w^d; I^d\right)}{\partial i} < 0$ for all i, w^d. Under this extension, integration has an additional benefit, represented by the third term on the right hand side of the following equation:

$$
\frac{\partial EV}{\partial i} = \frac{\partial v_1\left(s, i\right)}{\partial i} + \delta \int\limits_{w_R}^{\infty} \frac{\partial v^d\left(s, i\right)}{\partial i} f\left(w^d\right) dw^d + \int\limits_{w_R}^{\infty} v^d\left(s, i\right) \left(\frac{\partial f\left(w^d\right); I^d}{\partial i} \right) dw^d.
$$
[34]

Let us consider the relevant case in which migrants underestimate the benefits of integration on the wage distribution: $\frac{\partial \widetilde{F}\left(w^d; I^d\right)}{\partial i} > \frac{\partial F\left(w^d; I^d\right)}{\partial i}$. In this case, the misperception of the returns to integration detailed in **(1)** ($\frac{\partial E\widetilde{v}_2}{\partial i} - \frac{\partial E v_2}{\partial i}$) also contains an additional term, namely $\int\limits_{w_R}^{\infty} v^d\left(s, i\right) \left(\frac{\partial\left(\widetilde{f}\left(w^d; I^d\right) - f\left(w^d\right); I^d\right)}{\partial i} \right) dw^d$. Underestimating the benefits of integration on wages makes $\frac{\partial E\widetilde{v}_2}{\partial i} - \frac{\partial E v_2}{\partial i}$ even more negative, since the additional term is also negative, as one can show by integrating by parts:

$$
\begin{aligned}
\int\limits_{w_R}^{\infty} v^d\left(s, i\right) \left(\frac{\partial \left(\widetilde{f}\left(w^d; I^d\right) - f\left(w^d\right); I^d \right)}{\partial i} \right) dw^d &= \left[v^d\left(s, i\right) \left(\frac{\partial \widetilde{F}\left(w^d; I^d\right)}{\partial i} - \frac{\partial F\left(w^d\right); I^d}{\partial i} \right) \right]_{w_R}^{\infty} \\
&\quad - \int\limits_{w_R}^{\infty} \frac{\partial v^d}{\partial w^d} \left(\frac{\partial \left(\widetilde{F}\left(w^d; I^d\right) - F\left(w^d\right); I^d \right)}{\partial i} \right) dw^d.
\end{aligned}
$$

[34] As shown below, integrating by parts leads to the conclusion that $\int\limits_{w_R}^{\infty} v^d\left(s, i\right) \left(\frac{\partial f\left(w^d\right); I^d}{\partial i} \right) dw^d$ is always positive.

The first term on the right hand side is negative since

$$\left[v^d(s,i)\left(\frac{\partial\widetilde{F}\left(w^d;I^d\right)}{\partial i}-\frac{\partial F\left(w^d\right);I^d}{\partial i}\right)\right]_{w_R}^{\infty} = v^d(s,i)\left(\overbrace{\frac{\partial\widetilde{F}\left(\infty;I^d\right)}{\partial i}}^{=0}-\overbrace{\frac{\partial F\left(\infty\right);I^d}{\partial i}}^{=0}\right) - v^d(s,i)\left(\frac{\partial\widetilde{F}\left(w_R;I^d\right)}{\partial i}-\frac{\partial F\left(w_R\right);I^d}{\partial i}\right).$$

The second term on the right hand side is also negative since $\frac{\partial v^d}{\partial w^d}>0$ and $\frac{\partial\widetilde{F}\left(w^d;I^d\right)}{\partial i}>\frac{\partial F\left(w^d;I^d\right)}{\partial i}$. Summing up, when migrants underestimate the positive impact of integration on the distribution of wages, they (further) underestimate the marginal benefits to integration. Even if this is their only misperception, the reasoning applies identically, which also results in $\widetilde{i}<i^*$ and $\widetilde{s}>s^*$. □

A.4 Mispredicted long-term utility

A.4.1 Proof of Proposition 4

Proof. Let us start by reminding some identities. For any (s,i), $w_R(s,i)\equiv w_R\left(s,i;v^d,v^o\right)$ is such that $v^d(s,i;w_R(s,i))-v^o(s)=0$, and $\widetilde{w}_R(s,i)\equiv w_R\left(s,i;\widetilde{v}^d,\widetilde{v}^o\right)$ is such that $\widetilde{v}^d(s,i;\widetilde{w}_R(s,i))-\widetilde{v}^o(s)=0$.

Using these identities, let us first show that LT-pessimism implies $\widetilde{w}_R(s,i)\geq w_R(s,i)$ for any (s,i). Let us apply LT-pessimism for $w^d=w_R$: $\widetilde{v}^d(s,i;w_R(s,i))-\widetilde{v}^o(s)\leq v^d(s,i;w_R)-v^o(s)=0$. Since $\widetilde{v}^d(s,i;\widetilde{w}_R(s,i))-\widetilde{v}^o(s)$ is also equal to 0, LT-pessimism implies that for any (s,i), $\widetilde{v}^d(s,i;w_R(s,i))-\widetilde{v}^o(s)\leq\widetilde{v}^d(s,i;\widetilde{w}_R(s,i))-\widetilde{v}^o(s)$, or equivalently $\widetilde{v}^d(s,i;w_R(s,i))\leq\widetilde{v}^d(s,i;\widetilde{w}_R(s,i))$. This inequality is true if and only if $w_R(s,i)\leq\widetilde{w}_R(s,i)$.

Having proved that LT-pessimism implies $\widetilde{w}_R(s,i)\geq w_R(s,i)$, it follows naturally that LT-pessimism implies $\widetilde{p}=F\left(w_R\left(\tilde{s},\tilde{\imath};\widetilde{v}^d,\widetilde{v}^o\right)\right)\geq F\left(w_R\left(\tilde{s},\tilde{\imath};v^d,v^o\right)\right)=p_2$. This proves the sufficiency condition ($\Longrightarrow$).

Let us now prove the necessary condition, namely that if all migrants unexpectedly stay for a fixed misperception about long-term utility, then this misperception satisfies LT-pessimism. Formally, we need to prove that if $F\left(w_R\left(\tilde{s},\tilde{\imath};\widetilde{v}^d,\widetilde{v}^o\right)\right)\geq F\left(w_R\left(\tilde{s},\tilde{\imath};v^d,v^o\right)\right)$, then for all (v_1,F,z), $\widetilde{v}^d\left(s,i;w^d\right)-\widetilde{v}^o(s)\leq v^d\left(s,i;w^d\right)-v^o(s)$ for all (s,i). First note that $F\left(\widetilde{w}_R(\tilde{s},\tilde{\imath})\right)\geq F\left(w_R(\tilde{s},\tilde{\imath})\right)$ implies $\widetilde{w}_R(\tilde{s},\tilde{\imath})\geq w_R(\tilde{s},\tilde{\imath})$, and the latter inequality implies that $\widetilde{v}^d(s,i;\widetilde{w}_R(\tilde{s},\tilde{\imath}))\geq\widetilde{v}^d(s,i;w_R(\tilde{s},\tilde{\imath}))$. Thus, $\widetilde{w}_R(\tilde{s},\tilde{\imath})\geq w_R(\tilde{s},\tilde{\imath})$ implies $\widetilde{v}^d(\tilde{s},\tilde{\imath};\widetilde{w}_R(\tilde{s},\tilde{\imath}))-\widetilde{v}^o(\tilde{s})\geq\widetilde{v}^d(\tilde{s},\tilde{\imath};w_R(\tilde{s},\tilde{\imath}))-\widetilde{v}^o(\tilde{s})$, where the left hand side of the latter inequality is by definition equal to 0, and thus equal to $v^d(\tilde{s},\tilde{\imath};w_R(\tilde{s},\tilde{\imath}))-v^o(\tilde{s})$. Substituting this LHS, one obtains $v^d(\tilde{s},\tilde{\imath};w_R(\tilde{s},\tilde{\imath}))-v^o(\tilde{s})\geq\widetilde{v}^d(\tilde{s},\tilde{\imath};w_R(\tilde{s},\tilde{\imath}))-\widetilde{v}^o(\tilde{s})$. Since $(\tilde{s},\tilde{\imath})$ is defined for all possible (v_1,F,z), no restrictions are imposed on the possible values of $(\tilde{s},\tilde{\imath})$, so this condition must hold for any $(\tilde{s},\tilde{\imath})$. To conclude, simply note that this inequality for any $(\tilde{s},\tilde{\imath})$ is identical to the definition of LT-pessimism. □

A.4.2 Proof of Proposition 5

Proof. To prove that LTR-pessimism implies $\widetilde{s} > s^*$ and $\widetilde{i} < i^*$, it suffices to show that LTR-pessimism, that is $\partial\widetilde{u}^d/\partial c \leq \partial u^d/\partial c$ and $\partial\widetilde{u}^d/\partial I \leq \partial u^d/\partial I$ for all c and I, implies that $\frac{\partial E\widetilde{V}(s,i)}{\partial i} \leq \frac{\partial EV(s,i)}{\partial i}$ and $\frac{\partial E\widetilde{V}(s,i)}{\partial s} \geq \frac{\partial EV(s,i)}{\partial s}$ for all (s,i). One then obtains Proposition 5 by applying Proposition 1, similar to the proof of proposition 3. □

A.5 Mispredicted short-term utility

To prove the necessity of STpessimism for unexpected staying, we limit ourselves to biases that we call linear separable and monotonic. This restriction rules out situations where the bias in one dimension (savings or integration) depends on the utility in another dimension, as well as biases that sometimes over- and sometimes underestimate short-term returns.

Definition 8 (Linear separable and strictly monotonic bias)**.** *Let short-term utility $v_1(i,s)$ be separable: $v_1(i,s) = f(c_1(s)) + g(I) - k(i)$. Then a bias is linear separable if for the family of short-term preferences $v_1(i,s;\mu) = \mu \cdot f(c_1(s)) + g(I) - k(i)$ we have that $\tilde{v}_1(i,s;\mu) = \mu \cdot \tilde{f}(c_1(s)) + \tilde{g}(I) - \tilde{k}(i)$.*

A bias is strictly monotonic in c if either $\tilde{u}_1'(c) > u_1'(c)$ for all c or $\tilde{u}_1'(c) < u_1'(c)$ for all c.

A bias is strictly monotonic in i if either $\tilde{k}'(i) > k'(i)$ for all i or $\tilde{k}'(i) < k'(i)$ for all i.

A.5.1 Proof of Proposition 6

$\Longrightarrow$ By Proposition 7 (the proof of which does not depend on this proposition, so there is no circularity), we have that $\tilde{s} \geq \hat{s} = s^*$ and $\tilde{i} \leq \hat{i} = i^*$. Thus by Lemma 1, we have that $w_R\left(\widetilde{s},\widetilde{i}\right) \geq w_R\left(s^*,i^*\right) = w_R\left(\widehat{s},\widehat{i}\right)$, which implies that $\widetilde{p} \geq p_2$.

We will prove the following: *Suppose that we have a linear separable bias that is strictly monotonic in both i and c. Then a necessary condition for unexpected staying for all possible preferences migrant preferences is that $\tilde{k}'(i) \geq k'(i)$ and $\tilde{u}_1'(c) \leq u_1'(c)$.*

$\Longleftarrow$ Suppose the condition of the claim do not hold, so that there is some separable v_1 with either $\tilde{k}'(i) < k'(i)$ for all i or $\tilde{u}'(c) < u'(c)$ for all c.[35] Then it is enough to show that we can find a distribution of wages $F(\cdot)$ s.t. $\tilde{w}_R < w_R$ and such that intervals around these indifference wages are in the support of F – i.e. these migrants are unexpected *leavers*.

We will consider only migrant preferences of the following form: $v_1(s,i) = u(c_1(s)) + g(I) - k(i)$ that are separable, where $u(\cdot)$ is not purely the consumption utility in period 1, $g(\cdot)$ is the utility from integration, and $k(\cdot)$ as before is the cost of increasing integration. Morevoer, we assume that $v^o(s) = u(c^o(s)) + g(I^o)$ and $v^d(s,i,w) = u(c^d(s,w^d)) + g(I^d(i))$, that is, the consumption and integration utility are exactly the same. If we can show the condition is necessary for this restricted class of preferences, then the condition is certainly necessary for more general preferences.

[35] The proof can probably be generalized to allow for this condition to be satisfied only at some i and c.

Suppose the bias does not satisfy the condition for $v_1(s,i)$. Then consider the following family of preferences parameterized by μ:[36]

Proof. $v_1(s,i;\mu) \equiv \mu \cdot u(c(s)) + g(I) - k(i)$

- $v^o(s;\mu) \equiv \mu \cdot u(c^o(s)) + g(I^o)$
- $v^d(s,i,w;\mu) \equiv \mu \cdot u(c^d(s,w)) + g(I^d(i))$

Thus this family of preferences puts more and more weight on consumption utility as μ grows, and less as μ decreases.

First, note that the indifference wage at which the migrant is indifferent between returning and staying is well-defined even if it is not in the support of the wage distribution $F(\cdot)$ – in that case, it is simply the wage at which the migrant would be indifferent, *if* they were offered such a wage. Thus it is even well-defined for degenerate wage distributions that has a deterministic wage for sure.

Proof Overview The proof proceeds in three steps, each with substeps. First we consider a situation where we do not allow the migrants to return by assumption. We prove that in this simpler problem, fixing the wage distribution $F(\cdot)$, there is a μ such that $\tilde{w}_R^N(\tilde{s}, \tilde{i}, \mu, F) < w_R^N(\hat{s}, \hat{i}, \mu, F)$, where the superscript N highlights that there is "No return" – that this w_R is a different quantity, that captures at what wage the migrant would be indifferent between returning and staying if they were surprised in period 2 with this choice. Next, we create a wage distribution that always offers such high wages that all migrants with $\mu \geq \mu_L$ (for some well-calibrated value of μ_L) want to stay even for the lowest possible wage, no matter how much they save and integrate. Call this distribution F_0, which puts no weight on wages below $\bar{w}$. Since the migrant never expects to return for such a wage distribution, there problem is identical to the same wage distribution if they were not allowed to return. Hence by the first step, we can find a $\mu_0 = \mu(F_0)$ such that $\tilde{w}_R(\tilde{s}_0, \tilde{i}_0, \mu_0, F_0) = \tilde{w}_R^N(\tilde{s}_0, \tilde{i}_0, \mu_0, F_0) < w_R^N(\hat{s}_0, \hat{i}_0, \mu_0, F_0) = w_R(\hat{s}_0, \hat{i}_0, \mu_0, F_0)$. But even now, these anticipated and actual return wages are hypothetical and can never occur. Therefore we finally define F_ε as a wage distribution that puts total weight ε on wages below $\bar{w}$, distributed uniformly, and we show that as $\varepsilon \to 0$, $\tilde{w}_R(\tilde{s}(\varepsilon), \tilde{i}(\varepsilon), \mu_0, F_\varepsilon) \to \tilde{w}_R(\tilde{s}_0, \tilde{i}_0, \mu_0, F_0)$ and $w_R(F_\varepsilon)$ and similarly for $w_R(\varepsilon) \to w_R(0)$. This step will complete the proof, since it means that for some strictly positive ε we have that all wages for which the migrant would ever consider returning are possible (with incredibly low probability), hence the migrant will unexpectedly leave.

First, let us define a few quantities ex ante, to highlight that we can define them before starting to take limits, so that we don't end up with circular definitions.

No Return Allowed and Preliminaries Let us consider first the setup where migrants have no option of going home – but we can still ask for which (surprise) wage they *would* be indifferent

[36] It is easy to allow a different consumption utility in period 2, e.g. $u^d(c) = u^o(c) \neq u_1(c)$, but this is not necessary, since a counterexample from this more restricted family of preferences implies a counterexample for the larger family of preferences.

or anticipate being indifferent between returning and staying in period 2, based on their actual or anticipated savings and integration. We denote all the choice and anticipated variables for this altered problem with superscript N for "No Return" – so that $\tilde{s}^N$ are the anticipated savings under no return and $\hat{s}^N$ are the actual savings under no return. Then we have:

$$(\hat{i}^N(\mu, F), \hat{s}^N(\mu, F)) = \arg\max_{(i,s)} \mu \cdot u(S_0 - s) + (g(I) - k(i)) + \delta \int_0^\infty (\mu \cdot u(s + w) + g(I + i)) f(w)dw$$

$$(\tilde{i}^N(\mu, F), \tilde{s}^N(\mu, F)) = \arg\max_{(i,s)} \mu \cdot \tilde{u}(S_0 - s) + (\tilde{g}(I) - \tilde{k}(i)) + \delta \int_0^\infty (\mu \cdot u(s + w) + g(I + i)) f(w)dw$$

where S_0 is initial savings plus wage in the destination country in period 1. Here we used the fact that the bias is linear separable to be able to express the bias in a form that is linear in μ. Due to the separability of the utility function, there is no direct impact of the choice of i on s or vice versa, and because we dissallow return there is no indirect impact either through the actual return decision. Therefore we can rewrite this maximization over both (i, s) jointly as two independent maximizations over i and s:

$$\hat{s}^N(F) = \arg\max_s \mu \cdot u(S_0 - s) + g(I) + \delta \int_0^\infty \mu \cdot u(s + w)f(w)dw = \arg\max_s u(S_0 - s) + \delta \int_0^\infty u_d(s + w)f(w)dw$$

$$\hat{i}^N = \arg\max_i \delta \int_0^\infty g(I + i)f(w)dw - k(i) = \arg\max_i \delta g(I + i) - k(i)$$

and similarly for anticipated choices:

$$\tilde{s}^N(F) = \arg\max_s \tilde{u}(S_0 - s) + \delta \int_0^\infty u_d(s + w)f(w)dw$$

$$\tilde{i}^N = \arg\max_i \delta g(I + i) - \tilde{k}(i)$$

where none of the terms multiplied by δ are misperceived because they are about *future* utilities and outcomes which we assume are correctly perceived.

Notice that the arguments indicate that $\hat{i}^N$ and $\tilde{i}^N$ are independent of μ and F, while $\hat{s}^N$ and $\tilde{s}^N$ only depend on F but not on μ.

Before diving in, let us define a few quantities in order to highlight that they do not depend on our later choices of μ. If $\tilde{k}'(i) \geq k'(i)$ for all i, then set $\mu_L = 1$. If not, then let us show that since $\tilde{k}'(i) < k'(i)$ for all i (we assume that the bias distorts marginals in one direction), then we must have that $\tilde{i}^N > \hat{i}^N$. Suppose not, so that we have $\tilde{i}^N \leq \hat{i}^N$. Then (assuming FOCs yield the optimal choices) we have that

$$\delta g'(I + \hat{i}^N) \overbrace{=}^{FOC} k'^N(\hat{i}^N) \overbrace{>}^{\tilde{k}' < k'} \tilde{k}'^N(\hat{i}^N) \overbrace{\geq}^{\tilde{i}^N \leq \hat{i}^N} \tilde{k}'(\tilde{i}^N) \overbrace{=}^{FOC} \delta g'(I + \tilde{i}^N)$$

But $I + \hat{i}^N \geq I + \tilde{i}^N$, hence $g'(I + \hat{i}^N) \leq v'(I + \tilde{i}^N)$ by concavity of $g(\cdot)$, which contradicts the above. Hence $\hat{i}^N < \tilde{i}^N$.

Then we can define $\Delta(i,j) = g(I+i) - g(I+j)$ and $G(i) = g(I^o) - g(I+i)$, and pick any μ_L in $(0, \frac{\Delta(\tilde{i}^N, \hat{i}^N)}{u(x\cdot(\bar{s}+w_o))})$, which is a non-empty interval of strictly positive numbers since $\hat{i}^N < \tilde{i}^N$.

We can now define $\bar{w}$ which depends on μ_L as follows. Let $\bar{w}$ be the smallest wage for which the migrant with preferences determined by μ_L always stays, no matter how they save or integrate. Thus for every $w > \bar{w}$ the migrant stays even for $i = 0$ and $s = \bar{s}$ – such $\bar{w}$ exists, but we skip the proof as it is purely technical.

Now we define the wage distribution F_ε as uniform distribution on $[0, \bar{w}]$ with weight ε and uniform distribution on $[\bar{w}, 2\cdot\bar{w}]$ with weight $1-\varepsilon$. Formally:

$$f_\varepsilon(w) = \begin{cases} \frac{\varepsilon}{\bar{w}} \text{ for } w \leq \bar{w} \\ \frac{1-\varepsilon}{\bar{w}} \text{ for } w \in (\bar{w}, 2\cdot\bar{w}) \end{cases} \tag{A.1}$$

F_ε is constructed in such a way that both the biased and the unbiased migrant will always stay for $\varepsilon = 0$, i.e. for F_0. We now show that for any $\mu \geq \mu_L$, the migrant stays for any of the wages possible under F_0, since larger μ means they put more weight on money, which favors the destination country. By construction of F_ε, a migrant with μ_L-preferences stays for every wage $w > \bar{w}$:

$$\begin{aligned} &\mu_L \cdot u(s+w) + g(I+i) \geq \mu_L \cdot u(x\cdot(s+w_o^2)) + g(I^o) \\ \implies &\mu_L \cdot \left(u(s+w) - u(x\cdot(s+w_o^2))\right) \geq g(I^o) - g(I+i) \geq 0, \text{ since } I^o \geq I+i \end{aligned}$$

Hence $u(S+w) - u(x\cdot(S+w_o^2)) \geq 0$, so that for $\mu \geq \mu_L$ we have

$$\begin{aligned} \mu \cdot \left(u(s+w) - u(x\cdot(s+w_o^2))\right) &\geq \mu_L \cdot \left(u(s+w) - u(x\cdot(s+w_o^2))\right) \\ &\geq g(I^o) - g(I+i) \end{aligned}$$

which implies that the μ-migrant also prefers staying for this s and i – and since this holds for all s and i for μ_L, it holds for all s and i for $\mu \geq \mu_L$.

We now get to the main steps of the proof.

- **Step 1**: If $\tilde{u}'(c) > u'(c)$ then there is $\mu_0 \geq \mu_L$ s.t. $\tilde{w}_R^N(\mu_0, F_0) < w_R^N(\mu_0, F_0)$
- **Step 2**: If $\tilde{k}'(i) < k'(i)$ then there is $\mu \geq \mu_L$ s.t. $\tilde{w}_R^N(\mu_0, F_0) < w_R^N(\mu_0, F_0)$
- **Step 3**: If $\tilde{w}_R^N(\mu_0, F_0) < w_R^N$ for some μ_0, then there is $\varepsilon > 0$ s.t. $\tilde{w}_R(\mu_0, F_\varepsilon) < w_R(\mu_0, F_\varepsilon)$

Thus unless the conditions of the proposition are met, then by Steps 1 and 2 we have $\tilde{w}_R^N(\mu) < w_R^N(\mu)$ for some μ_0 when no return is possible. Hence by step 3, there is a wage distribution such that $\tilde{w}_R(\mu_0, F_\varepsilon) < w_R(\mu, F_\varepsilon)$, where return *is* possible, which proves the proposition.

Step 1 Let us show that if $\tilde{u}'(c) > u'(c)$ for all c, then $\hat{s}^N > \tilde{s}^N$ for all wage distributions, thus also for F_0. Suppose that $\hat{s}^N \leq \tilde{s}^N$ for some distribution of wages and let us derive a contradiction:

$$\hat{s}^N \leq \tilde{s}^N \implies S_0 - \hat{s}^N \geq S_0 - \tilde{s}^N \implies \tilde{u}'(S_0 - \hat{s}^N) \leq \tilde{u}'(S_0 - \tilde{s}^N) \tag{A.2}$$

by decreasing marginal utility. Thus, assuming FOCs characterize the unique interior maximand, we have

$$\delta \int_0^\infty u'^N(\hat{s}^N + w) f(w) dw \overbrace{=}^{FOC} u'(S_0 - \hat{s}^N) \overbrace{<}^{u' < \tilde{u}'} \tilde{u}'(S_0 - \hat{s}^N) \overbrace{\leq}^{\text{by A.2}} \tilde{u}'(S_0 - \tilde{s}^N) \overbrace{=}^{FOC} \delta \int_0^\infty u'(\tilde{s}^N + w) f(w) dw$$

This implies that $\tilde{s}^N + w < \hat{s}^N + w$, and hence that $\tilde{s}^N < \hat{s}^N$, directly contradicting our original assumption. Therefore $\tilde{s}^N < \hat{s}^N$.

We have the wage distribution F_0, for which as long as $\mu \geq \mu_L$ the migrant stays no matter which of the possible wages they receive, since all are above $\bar{w}$. Thus this choice is as if they couldn't return, hence we have $\hat{s}(F_0) = \hat{s}^N(F_0) > \tilde{s}^N(F_0) = \tilde{s}(F_0)$. Let us write $\tilde{w}_R(\mu) \equiv w_R(\tilde{s}(F_0), \tilde{i}, F_0, \mu)$ and $w_R(\mu) \equiv w_R(\hat{s}(F_0), \hat{i}, F_0, \mu)$ which are the indifference wages when the migrant acts when they can freely choose to return. Then for $\mu \geq \mu_L$:

$$\mu \cdot u(\tilde{s}^N + \tilde{w}_R^N(\mu)) + g(I + \tilde{i}^N) = \mu \cdot u(x \cdot (\tilde{s}^N + w_o)) + g(I^o)$$
$$\iff u(\tilde{s}^N + \tilde{w}_R^N(\mu)) = u(x \cdot (\tilde{s}^N + w_o)) + \frac{g(I^o) - g(I + \tilde{i}^N)}{\mu}$$

Thus, as $\mu \to \infty$, we have

$$\lim_{\mu \to \infty} \tilde{s}^N + \tilde{w}_R^N(\mu) = x \cdot (\tilde{s}^N + w_o)$$
$$\implies \lim_{\mu \to \infty} \tilde{w}_R^N(\mu) = (x - 1) \cdot \tilde{s}^N + x \cdot w_o$$
$$\implies \lim_{\mu \to \infty} w_R^N(\mu) - \tilde{w}_R^N(\mu) = (x - 1) \cdot (\hat{s}^N - \tilde{s}^N)$$

where the last line uses similar limits for w_R as for $\tilde{w}_R$. Since $\hat{s}^N > \tilde{s}^N$, this limit is larger than 0. Thus we can pick a sufficiently large μ_0 s.t. $\tilde{w}^N(\mu_0) < w^N(\mu_0)$. We are free to pick larger μ since F_0 has the property that migrants always stay if $\mu \geq \mu_L$. This proves step 1.

Step 2

If $\tilde{u}'(c) > u'(c)$ for all c, then we are done by applying step 1. Thus we can assume that $\tilde{u}'(c) \leq u'(c)$ for all c, so that $s^N \leq \tilde{s}^N$ for all wage distributions by the same argument as in step 1, but with the role of $\tilde{u}$ and u reversed. Moreover, as we proved in the preliminaries, we have that $\tilde{k}'(i) < k'(i)$ which implies, as we proved in the preliminaries, that $\hat{i}^N < \tilde{i}^N$.

When $\mu \to \mu_L$ from above, using the indifference conditions for $w_R(\mu)$ and $\tilde{w}_R(\mu)$ and using the fact that by construction $\mu_L < \frac{\Delta(\tilde{i}^N, \hat{i}^N)}{u(x\cdot(\bar{s}+w_o))}$ and $\Delta(\tilde{i}^N, \hat{i}^N) = g(I+\tilde{i}^N) - g(I+\hat{i}^N)$:

$$\begin{aligned}
\lim_{\mu\to\mu_L} \frac{u(\tilde{s}^N + \tilde{w}_R^N(\mu))}{u(\hat{s}^N + w_R^N(\mu))} &= \lim_{\mu\to\mu_L} \frac{\mu\cdot u(x\cdot(\tilde{s}^N + w_o)) + g(I^o) - g(I+\tilde{i}^N)}{\mu\cdot u(x\cdot(\hat{s}^N + w_o)) + g(I^o) - g(I+\hat{i}^N)} \\
&\leq \lim_{\mu\to\mu_L} \frac{\mu\cdot u(x\cdot(\bar{s} + w_o)) + g(I^o) - g(I+\tilde{i}^N)}{g(I^o) - g(I+\hat{i}^N)} \\
&= \frac{\mu_L\cdot u(x\cdot(\bar{s} + w_o)) + g(I^o) - g(I+\tilde{i}^N)}{g(I^o) - g(I+\hat{i}^N)} \\
&< \frac{\frac{\Delta(\tilde{i}^N, \hat{i}^N)}{u(x\cdot(\bar{s}+w_o))}\cdot u(x\cdot(\bar{s} + w_o)) + g(I^o) - g(I+\tilde{i}^N)}{g(I^o) - g(I+\hat{i}^N)} \\
&= \frac{\Delta(\tilde{i}^N, \hat{i}^N) + g(I^o) - g(I+\tilde{i}^N)}{g(I^o) - g(I+\hat{i}^N)} \\
&= 1
\end{aligned}$$

Hence there is μ_0 close enough to but larger than μ_L, we get

$$\begin{aligned}
\frac{u(\tilde{s}^N + \tilde{w}_R^N(\mu_0))}{u(\hat{s}^N + w_R^N(\mu_0))} < 1 &\iff u(\tilde{s}^N + \tilde{w}_R^N(\mu_0)) < u(\hat{s}^N + w_R^N(\mu_0)) \\
&\iff \tilde{s}^N + \tilde{w}_R^N(\mu_0) < \hat{s}^N + w_R^N(\mu_0) \\
&\iff \tilde{s}^N - \hat{s}^N + \tilde{w}_R^N(\mu_0) < w_R^N(\mu_0) \\
&\iff \tilde{w}_R^N(\mu_0) < w_R^N(\mu_0)
\end{aligned}$$

where the last line follows since $\tilde{s}^N - \hat{s}^N \geq 0$. This proves step 2.

Step 3

Since for F_0 the problem is as if migrants where not allowed to return, we have that the results hold for F_0 even when migrants are allowed, since wages are so high that returning is never optimal. Therefore we simply have to show that the problem is sufficiently well-behaved to apply Berge's theorem. Since all choice variables are by construction bounded, this can be done, but we skip it for brevity. □

A.5.2 Proof of Proposition 7

Proof. In the absence of other misperceptions, ST-pessimism is equivalent to $\frac{\partial E\widetilde{V}(s,i)}{\partial i} \leq \frac{\partial EV(s,i)}{\partial i}$ and $\frac{\partial E\widetilde{V}(s,i)}{\partial s} \geq \frac{\partial EV(s,i)}{\partial s}$ for all (s,i), hence by applying Proposition 1, one can directly show that ST-pessimism is sufficient for $\tilde{s} \geq s^*$ and $\tilde{\imath} \leq i^*$. Also note that in the absence of other misperceptions, $\left(\widehat{s}, \widehat{i}\right) = (s^*, i^*)$. Combining both informations, we have that $\tilde{s} \geq \widehat{s} = s^*$ and $\tilde{\imath} \leq \widehat{i} = i^*$ which is the result stated in Proposition 7. □

B Appendix: Data

B.1 Impact of pessimism on temporary intentions, remaining in Germany and unexpected staying

Table 4 provides results of linear regressions of the following equation:

$$Y_i = c + \beta Pessimism_i + \sum_j \gamma_j X_{ij} + e_i,$$

where each column corresponds to a different outcome variable Y_i related to unexpected staying: (1) a binary variable which equals 1 if the migrant stated temporary intentions at arrival and 0 otherwise, (2) a binary variable which equals 1 if the migrant is still in Germany in 2017 and 0 otherwise, and (3) our main measure of unexpected staying (UnexpS), i.e. the product of (1) and (2). The latter is thus a binary variable which takes value 1 if the migrant stays in Germany although they had expected to have left by 2017, and value 0 if they left Germany, or have correctly predicted to still be in Germany in 2017. In column (4), we use an alternative measure of unexpected staying: UnexpSL takes the value 1 in case of unexpected staying, 0 if initial intentions are aligned with actual location ex post, and -1 if the migrant has instead left unexpectedly. UnexpSL thus introduces more heterogeneity than the UnexpSL dummy and it captures the idea that unexpected staying and unexpected leaving are opposite phenomena. $Pessimism_i$ is measured in the first two years since arrival in Germany and pertains to either life satisfaction ($LS\_pessimism$) or wages ($wage\_pessimism$), and X_{ij} is a set of control variables capturing migrant characteristics.

These results show that migrants who were pessimistic about their wages did not form different intentions ex ante, but tend to stay more in Germany in the long term. As a result, the more migrants are pessimist about wages, the more they tend to stay unexpectedly. Having children is negatively associated with temporary intentions, and positively associated with actually staying in Germany ex post. Results are very stable with this alternative measure.

Table 5 provides results from similar regressions for pessimism about life satisfaction. Migrants who were pessimistic about their life satisfaction tend to formulate temporary intentions ex ante, but do not appear to have different actual return behaviors compared to more optimistic migrants. Since their intentions are more often temporary, these migrants are more likely to be unexpected stayers. Education and marriage are positively associated with staying ex post.[37] Tables 2 and 3 confirm that both types of pessimism are positively associated with unexpected staying under different sets of control variables. Changing the definition of unexpected staying (in column (4)) has a slight impact on significance of pessimism about life satisfaction, due to a minor decrease in the coefficient estimate combined with a slight increase in standard error, but overall results remain stable.

[37] We also controlled for having a partner abroad but no individual in the sample was in this situation.

Table 4: Intentions, final location, unexpected staying and wage-Pessimism

	(1)		(2)		(3)		(4)	
	Temp. intentions		Still in Germany		UnexpS		UnexpSL	
	b	se	b	se	b	se	b	se
Wage pessimism	0.034	(0.09)	0.164**	(0.08)	0.203**	(0.09)	0.199*	(0.11)
Age	-0.004	(0.01)	-0.010	(0.01)	-0.008	(0.01)	-0.013	(0.01)
Female	-0.100	(0.11)	0.094	(0.10)	-0.050	(0.11)	-0.006	(0.14)
Married	0.025	(0.13)	0.146	(0.11)	0.112	(0.12)	0.171	(0.15)
Children	0.115*	(0.07)	-0.113*	(0.06)	-0.053	(0.07)	0.002	(0.08)
Years of education	0.011	(0.02)	-0.015	(0.02)	-0.018	(0.02)	-0.004	(0.03)
Arrival cohort 1982-90	0.007	(0.18)	-0.237	(0.16)	-0.220	(0.17)	-0.230	(0.21)
Arrival cohort 91-2000	-0.205	(0.16)	-0.166	(0.14)	-0.343**	(0.16)	-0.371*	(0.19)
Constant	0.667*	(0.34)	1.271***	(0.30)	1.089***	(0.33)	0.939***	(0.41)
Observations	89		89		89		89	

Notes: "Temp. intentions"=1 if the immigrant states an initial intention to return and "Still in Germany"=1 if the migrant is still in Germany in 2017. "Unexpected stayers" (with=1) are migrants who report an initial intention to leave Germany but who stayed until 2017. ***,**,* denote significance at the 1, 5 and 10% level, respectively.

Table 5: Intentions, final location, unexpected staying and LS-Pessimism

	(1)		(2)		(3)		(4)	
	Temp. intentions		Still in Germany		UnexpS		UnexpSL	
	b	se	b	se	b	se	b	se
LS pessimism	0.049**	(0.02)	-0.011	(0.02)	0.041*	(0.02)	0.038	(0.03)
Age	-0.000	(0.00)	0.001	(0.00)	0.001	(0.00)	0.001	(0.00)
Female	-0.139	(0.08)	0.022	(0.06)	-0.089	(0.08)	-0.117	(0.09)
Married	-0.156	(0.12)	0.191**	(0.09)	-0.089	(0.11)	0.004	(0.02)
Children	-0.034	(0.05)	-0.063	(0.04)	-0.113**	(0.05)	0.036	(0.13)
Years of education	-0.025	(0.02)	0.029**	(0.01)	-0.012	(0.02)	-0.097	(0.06)
Arrival cohort 1982-90	0.209	(0.18)	0.155	(0.13)	0.357**	(0.17)	0.364*	(0.20)
Arrival cohort 91-2000	-0.012	(0.11)	-0.063	(0.08)	-0.017	(0.10)	-0.075	(0.12)
Constant	1.004***	(0.30)	0.368*	(0.22)	0.649**	(0.28)	0.372	(0.33)
Observations	141		141		141		141	

Notes: "Temp. intentions"=1 if the immigrant states an initial intention to return and "Still in Germany"=1 if the migrant is still in Germany in 2017. "Unexpected stayers" (with=1) are migrants who report an initial intention to leave Germany but who stayed until 2017. ***,**,* denote significance at the 1, 5 and 10% level, respectively.

Appendix B.2 shows how short-term decisions about integration and savings are linked to pessimism, temporary intentions and other covariates.

B.2 Impact of pessimism on short-term decisions (integration and savings)

In this section, we look at the endogenous decisions that are made in the short term (in the first two years after arrival in Germany). These decisions include integration and savings, which according to the model's predictions, should respectively decrease and increase with pessimism. Integration efforts i are proxied by the origin of newspapers that the migrant reads, while s is captured by the share of the migrant's monthly household income that is saved.

Table 6: Wage pessimism and short-term decisions

	(1)		(2)	
	Newspapers		saving share	
	b	se	b	se
Wage pessimism	-0.423	(0.35)	0.001	(0.02)
Temp. intentions	-0.332	(0.41)	-0.022	(0.03)
Age	0.048*	(0.03)	-0.001	(0.00)
Female	0.579	(0.45)	0.006	(0.03)
Married	-0.809	(0.50)	-0.051	(0.04)
Children	-0.136	(0.26)	0.012	(0.02)
Years of education	0.055	(0.10)	0.005	(0.01)
Arrival cohort 1982-1990	1.417*	(0.73)	0.002	(0.09)
Arrival cohort 1991-2000	1.009*	(0.52)	-0.043	(0.04)
Constant	0.204	(1.20)	0.140	(0.11)
Observations	32		60	

Notes: "Newspaper" is used as a proxy for integration and is defined as the origin of the newspapers read by the migrant. It is measured on a scale from 1 (only from the country of origin) to 5 (only from Germany). The saving rate is the proportion of savings in the household's monthly income. ***,**,* denote significance at the 1, 5 and 10% level, respectively.

Both measures of pessimism are negatively associated with integration as proxied by the type of journal that migrants read, though these estimates are not significantly different from 0. We do not find any link between pessimism and migrants' saving share.

Table 7: Pessimism about life satisfaction and short-term decisions

	(1)		(2)	
	Newspaper		saving share	
	b	se	b	se
LS pessimism	-0.110	(0.10)	-0.001	(0.01)
Temp. intentions	-0.156	(0.36)	0.023	(0.02)
Age	0.025	(0.02)	-0.000	(0.00)
Female	0.767*	(0.38)	0.005	(0.02)
Married	-0.637	(0.60)	-0.015	(0.02)
Children	-0.018	(0.22)	0.004	(0.01)
Years of education	0.165**	(0.08)	0.005	(0.00)
Arrival cohort 1982-1990	0.000	(.)	0.000	(.)
Arrival cohort 1991-2000	0.971**	(0.48)	-0.023	(0.02)
Constant	-1.248	(1.23)	0.023	(0.07)
Observations	52		119	

Notes: "Newspaper" is used as a proxy for integration and is defined as the origin of the newspapers read by the migrant. It is measured on a scale from 1 (only from the country of origin) to 5 (only from Germany). The saving rate is the proportion of savings in the household's monthly income. ***,**,* denote significance at the 1, 5 and 10% level, respectively.