Why do migrants stay unexpectedly? Misperceptions and implications for integration*

Marc Kaufmann† Joël Machado‡ Bertrand Verheyden§

July 5, 2023

Abstract

Empirical evidence suggests that many immigrants who initially plan to stay temporarily, eventually do not return to their country of origin. This unexpected staying suggests suboptimal integration among long-term migrants. We develop a theoretical framework in which migrants may misperceive their utility or wage prospects in the host country at arrival. Unaware of these misperceptions, they decide how much to integrate and save, incorrectly anticipating the effects of these actions on their location decision in the long term. We show that any migrant systematically underestimates their probability of staying if and only if their misperception involves specific forms of pessimism about the destination country relative to the country of origin, either in utility or in wage prospects. While present bias does not systematically lead to these forms of pessimism, biases such as projection bias, narrow bracketing, and misinference do – and thus can explain unexpected staying. Using the German Socio-Economic Panel, we find that higher levels of pessimism about utility and wages at arrival are associated with unexpected staying in the long term.

Keywords: Migrant integration ; return intentions ; unexpected staying ; misperceptions ; pessimism ; SOEP

JEL codes: F22 ; D91 ; J61

*The authors thank Andrea Albanese, Michel Beine, Michal Burzynski, Michael Clemens, 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, the 10th annual conference on Immigration in OECD countries, the 2021 EALE and LAGV annual meetings, the "Understanding voluntary and forced migration" conference, and the INTEGRIMM closing workshop 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.

†Central European University (CEU)

‡Luxembourg Institute of Socio-Economic Research (LISER)

§Luxembourg Institute of Socio-Economic Research (LISER)

1 Introduction

Over the last decades, growing migration flows have been at the heart of vivid societal debates. Socio-economic integration of immigrants is thereby seen as a key determinant of their impact on receiving countries. Integration efforts made by migrants are part of a complex process which involves social interactions, human capital investments, labor market participation and savings, among other aspects. All these dimensions are intertwined with migrants' long-term location plans (Dustmann and Görlach, 2016). In particular, the literature has shown that migrants with intentions to stay temporarily in the host country invest less in their integration than migrants with intentions to remain permanently (Dustmann, 1999; Geurts and Lubbers, 2017; Adda et al., 2022). This finding is important since temporary intentions at arrival are common. For instance, in their first appearance in the German Socio-Economic Panel (SOEP), about 37% of surveyed immigrants state that they plan to leave the country at some point in the future.[1]

However, return plans made in the early stages of migration often turn out to be incorrect. Substantial survey evidence shows that many migrants stay longer in the destination country than they initially intended, with some even staying permanently (see Schoorl, 2011 for the Netherlands, Adda et al., 2006; van Baalen and Muller, 2008; van den Berg and Weynandt, 2013 for Germany, Sinatti, 2011 and Agyeman, 2011 for Italy and Spain, Bolognani, 2007 for the UK, Alberts and Hazen, 2005 for the US, and Achenbach, 2017 for Japan). Table 1 draws on SOEP data to show that 77% of immigrants who expressed a temporary intention the first time they were surveyed had not left Germany as of 2020. We refer to this widespread underestimation of the probability of staying in the host country in the future as "unexpected staying".[2] Interestingly, the opposite pattern of unexpected leaving, i.e. immigrants initially stating permanent intentions but leaving, is far less frequent (i.e. less than 5% of migrants with permanent intentions, see Table 1).

Since integration is lower among immigrants with temporary intentions, and such intentions frequently do not materialize, the integration of many long-term immigrants is likely suboptimal. Understanding the causes and consequences of unexpected staying is thus critical to both immigrants and host countries.

In this paper, we investigate the role of behavioral mechanisms, or biases, in the formation of immigrants' perceptions about their prospects in the host country. We show that such mechanisms can explain unexpected staying. As migrants arrive in a country, they discover its customs, language,

[1] This proportion is based on SOEPv37, for all immigrants excluding refugees and ethnic Germans, surveyed since 1984 ; see Table 1. We show in Section 7.1 that refugees and ethnic Germans face different constraints and incentives, which materialize in stronger permanent intentions.

[2] Formally, unexpected staying does not occur if some migrants stay longer and some leave earlier. Unexpected staying requires that return intentions are systematically overestimated. In the empirical section, we explore alternative criteria to define unexpected staying, which lead to similar proportions. For instance, we exploit information on the intended duration of stay, which is however only available for a subsample of migrants who have temporary intentions.

institutions and labor market, whereas they weaken their direct links with their families and communities back home. This makes accurate predictions challenging, which might result in systematic mispredictions. While such misperceptions about the host country are natural and plausibly have important implications, they have received little attention in the literature.

Table 1: Migrants' initial return intentions and actual location in 2020

	Actual location in 2020		
Initial intentions	in Germany	left Germany	
Permanent stay	7,410	380	7,790
Temporary stay	3,537	1,047	4,584
	10,947	1,427	12,374

Note: SOEP data for all migrants (excluding refugees and ethnic Germans) surveyed between 1984 and 2020. Reported intentions to stay (temporarily or permanently) are collected from the immigrants' first reply to the relevant survey question.

We formally introduce misperceptions about migrants' future utility and long-term wage prospects in a model with endogenous integration, savings, and long-term return decisions. 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 plans to stay, the more they integrate and the less they save. Misperceptions at arrival about utility and wage prospects thus affect short-term decisions as well as the predicted probability of return. We determine the necessary and sufficient conditions for misperceptions such that they lead systematically to unexpected staying for all migrants – that is, independently of a migrant's observable and unobservable characteristics. To generate unexpected staying, these misperceptions must display specific forms of pessimism about the host country relative to the origin country, either in utility or in wage expectations. We then study behavioral mechanisms that can cause different misperceptions about wages, long-term utility and short-term utility.

We first study misperceptions about long-term wage prospects, which consists in a gap between the perceived distribution of wages and their actual distribution. We show that migrants have a tendency to stay unexpectedly and to integrate too little if and only if their perceived distribution is first-order stochastically dominated by the actual distribution. We argue that this misperception can result from incorrect inference. New cohorts of migrants may indeed base their long-term wage expectations on information provided by previous cohorts of migrants and return migrants. This information can however be biased, either because of negative selection or because of strategic misreporting. When failing to identify these biases, migrants tend to be pessimistic about long-term wages.

Next, we study misperceptions about long-term utility. We show that unexpected staying occurs for all migrants who are less optimistic about their well-being in case of stay in the destination

country than about their well-being in case of return to their origin country. This relative pessimism about the host country can arise from projection bias (Loewenstein et al., 2003) and narrow bracketing (Read et al., 1999). Projection-biased migrants underestimate how much they will adapt to the destination country, for instance because they anticipate 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 increasingly more integrated.[3]

To conclude the theoretical part, we study misperceptions about short-term utility. We find unexpected staying if and only if migrants have *asymmetric* misperceptions about the costs of their short-term actions (integration and savings). Migrants who think that integration will be demanding and that saving will be easy will initially have a strong intention to return. If later they realize that integration is simpler than they thought, whereas the opportunity cost of saving (enjoying consumption in the short term) is high, they are more likely to stay than they anticipated. This pattern cannot systematically be generated by present bias, which implies a *joint underestimation* of the costs of saving *and* integration. While present bias would imply unexpected staying in a model with exogenous integration, it does not systematically do so in our model.[4]

Second, we use SOEP data to explore the predictions of our model. We extract from this longitudinal dataset a sample of migrants who arrived in Germany between 1982 and 2010. Beyond relevant socio-demographic characteristics, the data contains information about migrants at different moments in time, such as their return intentions, beliefs about life satisfaction and wage prospects, their actual life satisfaction and wages in subsequent years, and whether they are still in Germany in 2020. We compare migrants' predictions to their actual outcomes, hence measuring migrants' mispredictions at arrival about their future location, life satisfaction and wage prospects. We measure pessimism about utility at arrival by comparing migrants' predicted life satisfaction to their realized life satisfaction five years later. A migrant is thus pessimistic about the host country if their level of predicted life satisfaction is lower than their actual level ex post.[5] Similarly, a migrant is pessimistic about wage prospects if they do not anticipate a wage increase which they actually obtain two years later. Controlling for migrant characteristics at arrival, cohort of arrival fixed effects, and life events occurring during the migration spell, we measure the association between

[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.

[4] Thus present bias can lead to unexpected staying for migrants for whom the savings effect dominates, and to unexpected leaving for migrants for whom the integration effect dominates. Its effect on unexpected staying is not systematic, but ambiguous.

[5] In line with Ivlevs (2015), we find that migrants are optimistic about their life satisfaction in the host country. This need not be inconsistent with the relative pessimism concept defined in our theory, which requires lower optimism about the destination country than about the origin country.

unexpected staying and pessimism at arrival. Consistent with the theory, we find that migrants who were more pessimistic about future life satisfaction and about wage prospects are significantly more likely to be unexpected stayers in 2020.

To sum up, our paper formally shows that migrants who are pessimistic about the host country overestimate their probability of return migration. This leads to socially inefficient outcomes by which a significant proportion of long-term migrants is suboptimally integrated. We discuss the sources of this pessimism through behavioral biases and misinformation prevalent in the early stages of migration. While public discourses tend to suggest that immigrants' poor socio-economic integration is due to an idealized short-term vision of the host country and thus to an excess of optimism, our results show that it is rather pessimism about the long term that leads migrants to stay longer than expected and integrate too little. Our results thus call for an improvement in the communication about policies reducing migrants' integration costs as well as detailed information about actual long-term prospects in both the destination and origin country.

We start by providing a literature review in Section 2. In Section 3, we introduce the theoretical framework. In Section 4, we first show that pessimism about long-term wages generates both unexpected staying and insufficient integration. We study misperceptions about utility in the long term in Section 5, and about utility in the short term in Section 6. In Section 7, we use SOEP data to empirically test the predictions of our model. We conclude in Section 8.

2 Related literature

Immigrants' unexpected staying has been documented in various developed countries. This phenomenon appears to have a systematic component, in the sense that it does not result from mere uncertainty or luck. 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 also present, though less pronounced, for Italian and Spanish migrants, who return to their home country in larger proportions. Furthermore, Steiner and Velling (1994) and van Baalen and Muller (2008) note that migrants' intended duration of stay keeps growing with the number of years spent in the destination country.

The literature has identified idiosyncratic shocks, such as life events, that can explain changes in return intentions (de Groot et al., 2011; Bettin et al., 2018). 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 their region of study if they accumulated location-specific capital. In addition, the socio-economic and political conditions in the origin country also affect immigrants' intended duration of stay (Kirdar, 2013).

However, unexpected shocks per se cannot explain the systematic pattern described above. Indeed, the structural inconsistency between intended and actual stay cannot result from "correct" beliefs about these shocks. 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.

An excess of optimism about labour market prospects could explain initial emigration decisions (Borjas and Bratsberg, 1996) but evidence in the literature is mixed. 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).

Our approach focusing on observed mispredictions (about wage and life satisfaction prospects) thus complements research explaining the heterogeneity in migrants' staying behavior and identifying idiosyncratic shocks that correlate with the likelihood of staying longer than intended. From this perspective, this study is the first to identify pessimism about the destination country as a general and systematic cause for migrants' unexpected staying. In particular, our analysis stresses that initial conditions at migrants' arrival have important impacts on long-term outcomes. This is in line with Fasani et al. (2021), who show that temporary employment bans, which cover the first months at arrival, reduce refugees' employment probability in the middle- to long-term by 15%.

Beyond improving the understanding of migrants' unexpected staying, 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 well documented 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; Adda et al., 2022), savings (Sinning, 2011), 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 well-being 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 naturalized, which is often considered as the final stage 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).

Finally, this paper contributes to the recent literature which integrates robust insights from behavioral economics into applied economics (Mullainathan et al., 2012; Eliaz and Spiegler, 2015; Chetty, 2015; Handel and Schwartzstein, 2018) in contrast to more common studies of separate behavioral models designed for specific biases. In line with this literature, our approach introduces misperceptions as a form of sufficient statistic for unexpected staying. This allows us to study multiple biases that systematically lead to pessimistic misperceptions within a general framework.

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. This setting is close to that of Adda et al. (2022), who develop a structural model of endogenous migrant integration, career paths and return migration. They show that migrants are affected by shocks which have persistent impacts on location preferences and lead to dynamic adjustments. Our focus is rather on explaining why migrants, at arrival, predominantly overestimate the probability that they will return in the future. To do so, we analyze potential behavioral mechanisms, or biases, that can lead to such systematic mispredictions. Our primary goal is indeed to explore the circumstances under which a migrant's return intentions are *mispredicted* in the early stage of migration. Our secondary goal is to explore

the implications this has for their migration and integration decisions.

Given these two goals, we develop a two-period model. In period 1 (the "short term"), the migrant has recently arrived in the destination country, and decides how much to integrate and save. These decisions are made in a context of uncertainty about their long-term labor market prospects. Also, beyond uncertainty, migrants may misperceive these prospects as well as their utility due to various behavioral mechanisms explained below. In period 2 (the "long term"), their prospects are revealed and they learn their true utility. Based on this information and on their past decisions, they decide whether to stay permanently 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. Migrants derive utility from consuming a numeraire good and from being integrated in the location where they currently live. The quantity of numeraire consumed for a given location L at a given period t is noted c_t^L, whereas the migrant's level of integration is noted I_t^L. In order to simplify notations, we drop location superscripts when they are not needed, i.e. in period 1 where the migrant is by definition in the destination country. In contrast, in period 2, two locations are possible, hence the superscript L is used, whereas the temporal subscript is not. Thus c_1 denotes consumption in the destination country in period 1, whereas I^L denotes the level of integration of the migrant in location L in period 2.

Modelling migrant integration Integration captures the sense of belonging and the ability to enjoy living in a certain location. The short-term level of integration in the destination country, I_1, is fixed at arrival as it relies on predetermined or exogenous factors, such as the degree of knowledge of the destination country's language and institutions before migrating, and the presence of family or ethnic networks at destination. In contrast, the long-term level of integration, I^d, can be increased by migrants' actions in the destination country. In period 1, migrants can for instance take language trainings, civic courses, assimilate the host country's customs or learn to appreciate its history and culture. These integration investments increase the period-2 level of integration in case of stay by an amount i, so that $I^d = I_1 + i$.[6] On the other hand, these investments require effort and impose in period 1 a utility cost noted $k(i)$, with $k'(i) > 0$, $k''(i) \geq 0$. In the baseline model, we assume that i and $k(i)$ are perfectly perceived by migrants. In Sections 5.2 and 5.3, we relax this assumption and study the impact of migrants' potential mispredictions of the relationship between integration investments and long-term utility.

In period 2, migrants' alternative to staying is to return to their origin country, where their integration level is I^o. We assume for simplicity that I^o is unaffected by i, which is destination-country specific, and that I^d cannot exceed I^o, i.e. that migrants cannot feel more integrated in the destination country than in their home country.[7] Misperceptions about I^o are discussed in Section

[6] In addition to the long-term level of integration, i can improve labor market outcomes. In Appendix A.4, we model this additional effect on the distribution of period-2 wages, on top of the direct effect of i on utility via I^d.

[7] The fact that I^o is unaffected by i is a simplification; what matters for our results is that i is relatively more

5.2.

Utility by period and location For c and I in any given period and location, utility is noted $u(c, I)$, with $\partial u(c, I)/\partial c > 0$, $\partial u(c, I)/\partial I > 0$, $\partial^2 u(c, I)/\partial c^2 \leq 0$, and $\partial^2 u(c, I)/\partial I^2 \leq 0$. Throughout the paper, we assume that $u(c, I)$ is separable ($\partial^2 u(c, I)/\partial I \partial c = 0$).[8] The only difference in the utility function between the two periods is the presence in period 1 of integration investment costs $k(i)$, so that period-1 utility is

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

where period-1 savings are deducted from consumption $c_1(s) = s_0 + w_1 - s$, and s_0 are exogenous savings accumulated before migrating and w_1 is the migrant's wage at arrival.

In the second period, utility depends on the migrant's location decision. Consumption in location L depends on the period-2 wage in location L, noted w^L. Given period-1 savings s, consumption in period 2 is thus either

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

where $x \geq 1$ is a real exchange rate which accounts for the higher purchasing power in the origin country compared to the destination country ($\partial c^o/\partial s = x \geq 1 = \partial c^d/\partial s$), and w^d and w^o are the migrant's wages in case of stay and of return, respectively.[9] Note that w^d is uncertain in period 1, as the migrant may not be aware of their long-term labor market prospects and potential wage assimilation. We first assume that the migrant only knows the distribution $F(w^d)$ and then study misperceptions regarding these long-term prospects in Section 4.

Since w^o and I^o are exogenous for a given migrant, the utility in case of return migration, $v^o(s)$, only depends on the period-1 savings decision s. In case of stay, utility $v^d(s, i; w^d)$, depends on s and integration efforts i, as well as on the realized wage w^d:

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

We now study the migrant's decisions regarding integration and savings in period 1, and return migration in period 2. Later, we introduce migrants' misperceptions about these elements and describe mechanisms leading to mispredictions about these decisions.

beneficial to I^d than to I^o. A sufficient condition for $I^d \leq I^o$ is $k'(I^o - I_1) \to \infty$.

[8] 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.

[9] As we explicit below, migrants may differ in their exogenous characteristics (s_0, I_1, x, I^o, w^o). We refer to this set of characteristics as a migrant "type".

Location decision: the reservation wage in the destination country First, we analyse the location choice, taking integration investments i and savings s as given. 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 $v^o\left(s\right)$. Return thus takes place if the realized wage in period 2 in the destination country is lower than a reservation wage w_R, which makes the migrant indifferent between both locations:

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

Note that the indifference wage is uniquely determined by the migrant's period-1 actions (s,i) and their period-2 preferences $(v^o\left(\cdot\right),v^d\left(\cdot\right))$, but does not directly depend on period-1 preferences $(v_1\left(\cdot\right))$, or on $F\left(\cdot\right)$. The threshold w_R together with the realized wage determines the migrant's decision to return or stay, and expectations over these two variables determine expected return decisions. Thus, a migrant who in period 1 perceives their long-term preferences to be $v^d\left(\cdot\right)$ and $v^o\left(\cdot\right)$, and who predicts period-1 actions as (s,i), anticipates correctly that they will return to the origin country with probability $F\left(w_R\left(s,i;v^d,v^o\right)\right)$. Any misprediction of w_R (through $v^d\left(\cdot\right)$ or $v^o\left(\cdot\right)$) or of $F\left(\cdot\right)$ will generate a misprediction about the probability of return, which will be the object of the following sections. Before that, we describe the optimal 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 their chosen (s,i) will lead to the period-2 reservation wage w_R, which also affects expected utility in period 2, noted Ev_2, with

$$Ev_2\left(s,i;v^d,v^o,F\right) = \int_0^{w_R} v^o(s)f(w^d)dw^d + \int_{w_R}^{\infty} v^d(s,i)f(w^d)dw^d.^{10} \tag{3.2}$$

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 $w_R = w_R(s,i,v^d,v^o)$. We 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 given a discount rate δ is to maximize $EV\left(s,i;v_1,v^d,v^o,F\right)$, with

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

where δ is a discount factor. The optimal period-1 actions (s^*,i^*) are given by:

$$(s^*,i^*) = \arg\max_{s,i} EV\left(s,i;v_1,v^d,v^o,F\right).^{11} \tag{3.3}$$

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

[11] We assume that there is a unique interior maximand for our maximization problem (see Figure 1 in the Appendix A.1).

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

$$p^* = F\left(w_R\left(s^*\left(v_1, v^d, v^o, F\right), i^*\left(v_1, v^d, v^o, F\right); v^d\left(\cdot\right), v^o\left(\cdot\right)\right)\right). \tag{3.4}$$

Migrants' predictions on their short term choices (s^*, i^*) and of their probability of return (p^*) depend on their perceptions. As highlighted by (3.3) and (3.4), (s^*, i^*) and p^*depend on migrants'$core parameters\left(v_1, v^d, v^o, F\right)$, that is, their preferences and their wage prospects. Clearly, if their perceptions about $\left(v_1, v^d, v^o, F\right)$are correct, migrants will correctly choose $(s^*, i^*)$and predict p^*. We now introduce misperceptions and describe their possible implications.

Introducing misperceptions about the migrants' core parameters Migrants may have misperceptins at arrival about both locations in the long term. These misperceptions apply to the migrant's *core parameters* $\left(\widetilde{v}_1, \widetilde{v}^d, \widetilde{v}^o, \widetilde{F}\right)$, which differ in at least one element from the actual core parameters $\left(v_1, v^d, v^o, F\right)$. We first show that such misperceptions affect short-term actions as well as the gap between predicted and actual probabilities of return migration in the long term. This allows us to formally define unexpected staying in the model, and to highlight the fact that misperceptions must satisfy specific conditions in order to generate this phenomenon.[12] In the subsequent sections, we then characterize these conditions for each type of misperception.

We now describe the migrant's maximization process under mispredictions made in period 1 about wage prospects and utility. First, taking (s, i)as given, migrants predict their probability of return in period 2, which may be impacted by mispredictions in wage prospects $\widetilde{F}$. The return decision depends on the reference wage threshold, which is also incorrectly predicted as soon as $\left(\widetilde{v}^d, \widetilde{v}^o\right)$is different from $\left(v^d, v^o\right)$. Migrants mispredicting future utility will indeed think that this threshold is $\widetilde{w}_R = w_R\left(s, i; \widetilde{v}^d, \widetilde{v}^o\right)$, which is such that $\widetilde{v}^d\left(s, i; \widetilde{w}_R\right) = \widetilde{v}^o\left(s\right)$. These distorsions of the predicted probability of return also impact period-1 decisions, since these are based on the tradeoff between short-term costs and (incorrectly predicted) future expected benefits:

$$(\widetilde{s}, \widetilde{i}) = \arg\max_{s,i} E\widetilde{V}, \tag{3.5}$$

where $E\widetilde{V} = EV\left(s, i; \widetilde{v}_1, \widetilde{v}^d, \widetilde{v}^o, \widetilde{F}\right)$. For these actions $(\widetilde{s}, \widetilde{i})$, the mispredicted probability of return is noted:

$$\widetilde{p} = \widetilde{F}\left(w_R\left(\widetilde{s}, \widetilde{i}; \widetilde{v}^d, \widetilde{v}^o\right)\right). \tag{3.6}$$

Summing up, mispredictions $\left(\widetilde{v}^d, \widetilde{v}^o, \widetilde{F}\right)$ distort both period-1 actions $(\widetilde{s}, \widetilde{i})$ and the predicted probability of return $\widetilde{F}\left(\widetilde{w}_R(s, i)\right)$. When a migrant *over*estimates the probability that they will return,

[12] Some forms of misperceptions might instead generate unexpected leaving, whereas other misperceptions might generate one or the other depending on migrants' types.

unexpected staying pertains to structural errors in this predicted probability of return $\widetilde{p}$. This overestimation occurs if $\widetilde{p}$ is larger than p_2, the *actual* return probability for this choice $(\widetilde{s}, \widetilde{i})$. This actual probability p_2 relies on (i) past actions $(\widetilde{s}, \widetilde{i})$, (ii) the actual wage distribution F, and (iii) on the fact that, in period 2, migrants know their actual preferences v^d, v^o:

$$p_2 = F\left(w^R\left(\widetilde{s}, \widetilde{i}; v^d, v^o\right)\right). \tag{3.7}$$

Definition 1. Unexpected staying *means that the migrant overestimates their probability of return migration:* $\widetilde{p} > p_2$.

Throughout the rest of the paper, we seek to identify conditions about misperceptions which are both sufficient and necessary to generate an overestimation of the probability of return. Most importantly, these conditions must generate unexpected staying *for all possible migrant characteristics.*[13] We refer to this set of characteristics as a migrant's multidimensional "type".

Definition 2. *A migrant type* $z = (s_0, I_1, I^o, w^o, x)$ *is a set of predetermined characteristics that migrants correctly perceive.*

Our focus on conditions that generate unexpected staying for all migrant types stems from the fact that if a migrant's misperceptions satisfy such conditions, then they must overestimate their return probability even if we know nothing about their type z.[14] If the conditions on misperceptions were not sufficient for all z, this would mean that these misperceptions might lead some migrants to stay unexpectedly while some other migrants might *leave* unexpectedly. Conversely, if the conditions were not necessary for all z, we might wrongly attribute unexpected staying to these conditions, whereas unexpected staying could result from the specific characteristics of a given migrant. While necessary and sufficient conditions are naturally stronger, they allow us to weaken the data requirements to test our predictions. These requirements are already high as we illustrate in Section 7. Hence, we do not cover misperceptions that may lead to unexpected staying for some z, but rather focus on those that lead to unexpected staying for all z.

To make clear statements about unexpected staying, we separately review (i) misperceptions about long-term utility (the shape of the utility functions in the long term $v^d(\cdot,\cdot)$ and $v^o(\cdot)$), (ii) misperceptions about the distribution of labor market outcomes (the cumulative density function $F\left(w^d\right)$), and (iii) misperceptions about short-term utility $v_1(\cdot,\cdot)$. For each of these misperceptions, we describe the conditions that they should satisfy to generate unexpected staying for all z, and discuss mechanisms that may generate these misperceptions.

[13] Our theory does not impose any restrictions on migrants' characteristics, except for the fact that they should be allowed to work in the destination country, and that they should be able to decide their location in the long term.

[14] For example, we can make statements about unexpected staying even if we do not know their integration level at arrival (I^1) or at origin (I^o), nor their wage in the origin country (w^o).

4 Misperceiving probabilities about wage prospects

In this section, we focus on a migrant with incorrect beliefs about *probabilities*, which means their 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: $\left(\widetilde{v}_1, \widetilde{v}^d, \widetilde{v}^o\right) = \left(v_1, v^d, v^o\right)$.

As shown in equation (3.1), the reservation wage w_R is independent of $\widetilde{F}(\cdot)$, and thus in this Section the migrant correctly predicts w_R. However, migrants mispredict their probability of return, as a direct consequence of the difference between $\widetilde{F}(w_R)$ and $F(w_R)$.

In Subsection 4.1, we define which form of pessimism about wages is necessary and sufficient to lead to unexpected staying. In Subsection 4.2, we discuss how migrants may develop such pessimistic wage expectations.

4.1 Pessimistic misperceptions of wage prospects

We first provide a definition of misperceptions about long-term host-country wage prospects $\widetilde{F}(\cdot)$ which display pessimism.

Definition 3. *A migrant is assumed to be* ***pessimistic about wage prospects*** *in the destination country (**W-pessimistic**) if the misperceived wage distribution $\widetilde{F}(\cdot)$ is first-order stochastically dominated by the actual distribution $F(\cdot)$, i.e. $\widetilde{F}(w^d) > F(w^d)$ for all w^d.*

A W-pessimistic migrant expects lower wages than they are likely to receive in the destination country.[15] Proposition 1 states that this misperception is necessary and sufficient to produce unexpected staying in a systematic way.

Proposition 1. *Consider all possible migrant types $\left(z, v_1, v^d, v^o\right)$ with the same misperception $\widetilde{F}(\cdot)$. Unexpected staying occurs for all $\left(z, v_1, v^d, v^o\right)$ if and only if their misperception $\widetilde{F}(\cdot)$ satisfies W-pessimism.*

Proof. See Appendix. □

W-pessimism is a sufficient condition for migrants to overestimate the probability that they will return: $\widetilde{F}(w_R) > F(w_R)$ for any w_R. W-pessimism is on the other hand necessary to ensure that unexpected staying occurs for all migrants, i.e. all migrants with (i) any possible values of (correctly perceived) core parameters $\left(v_1, v^d, v^o\right)$ and any possible type $z = (s_0, I_1, x, I^o, w^o)$, and (ii) whose only common feature is their W-pessimistic misperception $\widetilde{F}(\cdot)$. To fix ideas about why W-pessimism is necessary, consider a misperception $\widetilde{\widetilde{F}}(\cdot)$ which is not first-order stochastically dominated by $F(\cdot)$: $\widetilde{\widetilde{F}}(\cdot)$ crosses $F(\cdot)$ once at some wage $\widetilde{\widetilde{w}}^d$, such that $\widetilde{\widetilde{F}}\left(w^d\right) > F\left(w^d\right)$ for

[15] 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.

$w^d < \widetilde{\widetilde{w}}^d$, and $\widetilde{\widetilde{F}}\left(w^d\right) < F\left(w^d\right)$ for $w^d > \widetilde{\widetilde{w}}^d$. While $\widetilde{\widetilde{F}}(\cdot)$ can generate unexpected staying for migrants $\left(z, v_1, v^d, v^o\right)$ for whom $w_R\left(z, v_1, v^d, v^o\right) < \widetilde{\widetilde{w}}^d$, it generates instead unexpected *leaving* for all other migrants for whom $w_R\left(z', v_1', v^{d\prime}, v^{o\prime}\right) > \widetilde{\widetilde{w}}^d$.

Beyond mispredicting return probabilities, misperceptions about $F(\cdot)$ impact migrants' short-term actions. In the case of W-pessimism, migrants save too much and integrate too little compared to what would have been optimal for them: $\widetilde{i} < i^*$ and $\widetilde{s} > s^*$.[16] Since savings bring higher utility in the origin country, the W-pessimistic migrant, who overestimates the probability that they will return, overestimates the incentives to save. Conversely, they integrate too little because they underestimate the probability that they will enjoy the benefits of integration.[17]

4.2 Misinference from migrants as a cause for wage pessimism

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) and Seshan and Zubrickas (2017) document that potential earnings from migration tend to be underestimated in the origin country. On the other hand, Shrestha (2020) finds that potential 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.

When forming expectations about their own prospects, prospective migrants may base these expectations on the labor market outcomes in the destination country reported by other migrants. Such expectations may be biased for three reasons. First, return migrants tend to be negatively selected on average among the whole population of migrants (Abramitzky et al., 2019). 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. Second, 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). Migrants may form pessimistic anticipations because they fail to account for the strategic incentives in immigrants' reporting. Various models have been proposed in which agents misinfer in such strategic settings, along with lab evidence of neglect of such strategic concerns and selection issues (Eyster and Rabin, 2005; Esponda, 2008; Esponda and Vespa, 2014). Third, many

[16] We provide a formal treatment of this question in Appendix A.3, but present here the main intuition.

[17] As a result, note that despite generating unexpected staying, W-pessimism affects migrants' actions in such a way that they return more than if they had saved and integrated optimally.

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

migrants first settle in ethnic enclaves, which facilitate their initial steps in the destination country. However, migrants living in these enclaves can be negatively selected if successfully integrated migrants are more likely to leave them, which can lead recently arrived immigrants to misinfer their wage prospects.

5 Misperceiving long-term preferences

In this section, we consider errors that are exclusively due to the migrant misperceiving utility in the long term (i.e. in period 2). While their actual preferences ex post are $v^d(\cdot,\cdot;\cdot)$ and $v^o(\cdot)$, their predictions of these preferences in period 1 are $\widetilde{v}^d(\cdot,\cdot;\cdot) \equiv \widetilde{u}^d\left(c^d, I^d\right)$ and $\widetilde{v}^o(\cdot) \equiv \widetilde{u}^o(c^o, I^o)$.

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

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

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

By maximizing misperceived expected utility $E\widetilde{V}$, migrants predict in period 1 that their integration and savings are $(\widetilde{s},\widetilde{i})$, which makes them think that they will return to the country of origin if long-term wages w^d are below $\widetilde{w}_R(\widetilde{s},\widetilde{i})$. Their predicted probability of return is thus $\widetilde{p} = F\left(\widetilde{w}_R(\widetilde{s},\widetilde{i})\right)$. In contrast, at the start of period 2, they realize that with their actual preferences (v^d, v^o), their reservation wage is in fact $w_R(\widetilde{s},\widetilde{i}) = w_R\left(\widetilde{s},\widetilde{i};v^d,v^o\right)$. Their actual probability of return is thus $p_2 = F\left(w_R(\widetilde{s},\widetilde{i})\right)$. We now study the conditions under which misperceptions $(\widetilde{v}^d,\widetilde{v}^o)$ systematically generate unexpected staying $(\widetilde{p} > p_2)$.

5.1 Relative pessimism about long-term utility in the host country

We define misperceptions $(\widetilde{v}^d,\widetilde{v}^o)$ which display relative pessimism about the destination country.

Definition 4. *A migrant is assumed to be* ***relatively pessimistic about their long-term utility in the host country*** *compared to the origin country (**LT-pessimistic**) if* $\widetilde{v}^d\left(s,i;w^d\right) - \widetilde{v}^o(s) < v^d\left(s,i;w^d\right) - v^o(s)$ *for any given* i*,* s*, and* w^d*.*

Pessimism is relative in the sense that misperceptions about the destination country should be compared to misperceptions about the origin country in the long term. In fact, migrants may be relatively pessimistic even though they are *optimistic* about the destination country ($\widetilde{v}^d\left(s,i;w^d\right) - v^d\left(s,i;w^d\right) > 0$). This is indeed the case if they overestimate long-term utility in the origin country even more, i.e. if

$$\widetilde{v}^o(s) - v^o(s) > \widetilde{v}^d\left(s,i;w^d\right) - v^d\left(s,i;w^d\right) > 0.$$

This possibility is relevant for different reasons. Migrants may underestimate the fact that their reference points in terms of standards of living, expectations about amenities or health services have changed. In addition, they might overestimate the resilience of their social capital in the origin country or even the origin country's socio-economic development. The following proposition formalizes that this relative pessimism is necessary and sufficient for unexpected staying.

Proposition 2. *Consider all possible migrants* (z, v_1, F) *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* $\left(\widetilde{v}^d, \widetilde{v}^o\right)$ *satisfies LT-pessimism.*

Proof. See Appendix A.5. □

The intuition behind Proposition 2 is that since LT-pessimistic migrants underestimate utility in the destination country relative to the origin country, they overestimate the long-term wage they will require to stay in the destination country: $\widetilde{w}_R\left(\widetilde{s}, \widetilde{\imath}\right) > w_R\left(\widetilde{s}, \widetilde{\imath}\right)$. Consequently, even though they correctly perceive their labor market prospects ($F(\cdot)$), LT-pessimism makes the migrant overestimate their probability of return migration: $\widetilde{p} = F\left(\widetilde{w}_R\right) > F\left(w_R\right) = p_2$.

This result implies that when migrants are pessimistic about long-term utility, unexpected staying does not occur under any circumstances; instead it occurs for specific, intermediate, realizations of long-term wages. More specifically, for all wages w^d below the actual reservation wage w_R, migrants correctly predict that they will return. Conversely, for all wages above their misperceived reservation wage ($w^d > \widetilde{w}_R$), migrants correctly predict that they will stay. It is only for intermediate values of w^d (between w_R and $\widetilde{w}_R$) that migrants' predictions are incorrect. Indeed, for wages $w^d \in \left(w_R; \widetilde{w}_R\right)$ migrants predict that they will return (since $w^d < \widetilde{w}_R$) whereas ex post they stay (since $w^d > w_R$). Unexpected staying thus occurs when wage realizations belong to this interval, which has a probability equal to $F\left(\widetilde{w}_R\right) - F\left(w_R\right)$.

The next question we consider is how an LT-pessimistic migrant actually behaves in terms of savings and integration in period 1 $\left(\widetilde{s}, \widetilde{\imath}\right)$ compared to a migrant whose utility is correctly predicted (s^*, i^*). While one might expect that LT-pessimism systematically leads to insufficient integration and excessive savings, this is not always the case. LT-pessimistic migrants may have integrated more than the optimal level and yet overestimate their probability of return.[19] Thus, a stronger condition is needed, based on a comparison of marginal utilities instead of a comparison of utility levels: to integrate too little and save too much, migrants should underestimate the relative marginal benefits of integration and savings in the host country.[20] In the next section, we study behavioral biases which can cause LT-pessimism, and show that when migrants are projection-biased, they tend to stay unexpectedly *and* to integrate too little.

[19] However, this counterintuitive scenario need not occur systematically.

[20] We provide a more thorough description of this result is in Appendix A.5.

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 migration in the German SOEP. Projection bias refers to the tendency of people to perceive their future tastes as more similar to their current tastes than they actually will be.[21] For example, migrants might initially dislike the food, the weather and the customs, all of which are different from their origin country. Limited proficiency in the local language also contributes to make them feel alienated. These migrants thus suffer at arrival from a low level of integration I_1 compared to the integration level they had in the country that they just left ($I^o > I_1$). Under projection bias, the migrant perceives their future utility in the destination country as lying between the actual period-2 utility (with I^d) and the utility this would yield in period 1 (with I_1). Therefore, since $I_1 < I^d$, the projection-biased migrant underestimates their utility $u\left(c^d, I^d\right) = v^d(s, i; w^d)$ for all $\left(s, i; w^d\right)$.[22]

Note that, conversely, it could be assumed that over the course of period 1 the migrant loses some of their initial integration in the origin country I^o. For instance, I^o could decrease because migrants lose some of their social ties in the origin country. Projection bias about the origin country would in this case imply that in period 1 the migrant overestimates the utility that they would derive in case of return in period 2. Eventually, both of these effects (underestimation of v^d and overestimation of v^o) go in the same direction and are consistent with LT-pessimism.

Formally, we apply the model of simple projection bias from Loewenstein et al. (2003) to our framework. We focus here on misperceptions of utility in case of stay, which is noted:

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

where $\widetilde{u}(c^d, I^d | I_1, \alpha)$ denotes the period-2 utility as perceived in period 1, 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.

Projection bias leads to long-term pessimism, since misperceived utility is lower than actual utility, for all consumption and integration levels:

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

[21] 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, 2019; Bushong and Gagnon-Bartsch, 2020), as well as for people mispredicting their habit formation for gym attendance.

[22] Unlike Dustmann and Görlach (2016); Adda et al. (2022) 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 present in Appendix A.4 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.

since for any given c^d, $u(c^d, I^d) > u(c^d, I_1)$ by $\partial u/\partial I > 0$. Thus, we have that for $\alpha > 0$, $\widetilde{v}^d\left(s, i; w^d\right) - v^d\left(s, i; w^d\right) \leq \widetilde{v}^o(s) - v^o(s)$, which is the definition of LT-pessimism. Note that we only assumed in this formalization of projection bias that it affects v^d, so that $\widetilde{v}^o(s) = v^o(s)$. However, if we were to also consider the possible impact of projection bias on the origin country, it would imply that $\widetilde{v}^o(s) > v^o(s)$, which amplifies the degree of LT-pessimism. Hence, applying Proposition 2, we can conclude that projection bias leads to unexpected staying.

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).[23] We now model a form of narrow bracketing that we call incidental integration, and which generates unexpected staying but does not systematically lead to non-optimal integration. Suppose that the integration process can be divided into two categories, namely active efforts such as language trainings (i in our baseline model) and basic day-to-day interactions with natives j, so that

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

While both types of actions contribute to I^d, only i is perceived by the migrant as a conscious investment in view of increasing I^d. In other words, j incidentally increase the level of long-term integration. Formally, j can be seen as a mechanical, exogenous, action by the migrant, or it could be modeled as an additional decision variable whose positive impact is perceived only in period 1 by the migrant.[24] The key point here is that a migrant who narrowly brackets j from i will ignore the effect of j when they predict their long-term integration level:

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

where $0 < \gamma \leq 1$ defines the fraction of j ignored by the migrant. 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, there is no incidental integration in the origin country ($\widetilde{v}^o = v^o$). As a result, narrow bracketing implies LT-pessimism and leads to unexpected staying. However, the migrant does not underestimate the returns to integration and hence we cannot conclude that they always integrate less and save more.[25]

[23] 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).

[24] We could model it as: $v_1(s, i, j) = u(c_1(s), I_1(j)) - k(i) - l(j)$.

[25] See Appendix A.5 for a more detailed discussion about insufficient integration.

6 Mispredicting short-term preferences

We finally turn our focus to models in which the migrant mispredicts their utility in period 1. To this end, we introduce in this section a period 0. In this preliminary period, the migrant has just arrived in the destination country and predicts their period-1 actions (savings and integration). Thus in period 0, the migrant may misperceive v_1 as $\widetilde{v}_1$.[26] For instance, the migrant may in period 0 underestimate the cost of integration $k(i)$, or the opportunity cost of savings (that is, the marginal utility of period-1 consumption). For example, the migrant in period 0 may plan to save a lot in period 1 (in view of preparing a return in period 2), but eventually does not manage to accumulate such target savings because renouncing consumption turned out to be more demanding than expected. Formally, they predict in period 0 their choices to be $(\widetilde{s}, \widetilde{\imath})$ based on maximizing:

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

where the correctly perceived period-2 expected utility $Ev_2\left(s, i; v^d, v^o, F\right)$ is defined as in Equation (3.2). Since $\widetilde{v}_1$ differs from v_1, predictions made in period 0 about $(\widetilde{s}, \widetilde{\imath})$ and in turn about the probability of return $\widetilde{p} = F\left(w_R(\widetilde{s}, \widetilde{\imath})\right)$ are incorrect. In period 1, migrants realize that their actual preference are v_1, and thus choose -in the absence of other misperceptions- the optimal actions (s^*, i^*), since their period-1 objective is the same as the one defined in 3.3. Hence, in period 1, they correctly predict their return probability to be $p_2 = p^* = F\left(w_R\left(s^*, i^*\right)\right)$. Unexpected staying under present bias thus occurs if $\widetilde{p} > p^*$.

As in previous sections, we start by studying the conditions on $\widetilde{v}_1$ which are necessary and sufficient to generate unexpected staying for all migrant types, and then study whether specific bias can satisfy these conditions. We provide here a summary of the results and leave their full exposition to the Appendix A.6.

First note that unexpected staying ($\widetilde{p} > p^*$) occurs if and only if mispredicted period-1 actions $(\widetilde{s}, \widetilde{\imath})$ in period 0 lead to a mispredicted reservation wage $w_R(\widetilde{s}, \widetilde{\imath})$ that is higher than the actual $w_R\left(s^*, i^*\right)$. This occurs if the migrant's mispredicted actions lead them to think that they will enjoy the destination country in period 2 less than they actually will. This is the case if in period 0 they think they will integrate less than they actually will, and think they will save more than they actually will. For this to occur, migrants' misperceived $\widetilde{v}_1$ in period 0 would have to (i) *under*estimate the cost of saving from period 1 to period 2, and (ii) *over*estimate the cost of integration in period 1. In period 1, however, when they realize that saving is actually more costly and that integration is less costly than they thought, they save less, integrate more, and end up staying more frequently than they had predicted.

We then study whether biases about preferences in the short term can satisfy this asymmetric pair of conditions on $\widetilde{v}_1$. The most common type of bias about short-term preferences is naive

[26] We assume that other core parameters $\left(v^d, v^o, F\right)$ are correctly perceived.

present bias. However, present bias is not compatible with the requirement that mispredictions over the costs of integration and savings go in opposite directions. Indeed, present bias implies a hyperbolic discounting of utility in future periods. Since this discounting applies to the whole utility function at a given period, the present-biased migrant *uniformly* underestimates in period 0 the costs of *both period-1 actions*, namely *savings and integration*. This violates the conditions described above, as they require instead an *asymmetric* misperception of the costs of these two actions.

Returning to target savings, present bias leads migrants to save less than they had intended, which implies that they would be less likely to return than initially planned. However, since the costs of both actions were underestimated, present bias also leads them to integrate less than they thought. Hence, in a model with endogenous integration, it is not clear whether these migrants are less, or more, likely to return than initially planned.

Finally, note that this does not mean that present bias can never generate unexpected staying. Instead, it shows that there are migrant types for which present bias implies unexpected leaving rather than unexpected staying.

7 Stylized facts from German SOEP

In this section, we aim to document whether, consistent with our theory, initial pessimistic misperceptions (about utility or wage prospects) are associated with an increased likelihood of staying unexpectedly in Germany. In Section 7.1, we first describe the German Socio-Economic Panel (SOEP) and the sample of migrants on which our analysis is based. We further discuss the demanding assumptions required to bring the theory as close as possible to the data. In Section 7.2, we provide descriptive statistics and correlations between unexpected staying and two forms of pessimism: wage-pessimism and pessimism about life satisfaction.

7.1 Data

The data and sample criteria The SOEP is a survey which provides household- and individual-level data for a representative sample of the population in Germany on a yearly basis over several decades (Goebel et al., 2019).[27] This population includes immigrants observed since the first wave in 1984 to the last at our disposal in 2020. In addition to standard demographic and socio-economic characteristics, the panel includes information about immigrants' country of birth, year of arrival and last known location (i.e. whether the migrant left Germany). To the best of our knowledge, it is the only long panel containing the information required for our analysis on the following key variables: immigrants' initial intentions (to stay permanently or temporarily), proxies for initial

[27] Version 37, SOEP, 2022, doi: 10.5684/soep.core.v37eu.

levels of pessimism and for short-term levels of integration, savings, and actual location in the long term.

Our sample of migrants does not impose restrictions on the type of intentions they had at arrival, nor on their latest location. Hence, our sample consists of migrants who had either temporary or permanent initial intentions, and who have either left Germany or stayed. Migrants can thus be unexpected stayers, unexpected leavers, or have made correct predictions about the temporariness or the permanent nature of their stay.

However, since the focus of our paper is to explain why immigrants' *early* return intentions are frequently wrong, and how this impacts their behavior in the short term, we need to ensure that this information is collected sufficiently close to the time of arrival. Also, long-term return behaviors should be measured sufficiently long after arrival in order to limit censoring issues. Hence, we need to impose a number of selection criteria to our sample, which are explained in detail below. Table A1 describes the incremental impacts of these selection criteria on sample sizes, more specifically on the number of immigrants for which we observe LS-pessimism (column 1) and W-pessimism (column 2) respectively.

First, the variables required to build our pessimsim measures are not available for all respondents. Our first criterion is thus the availability of respondents' predictions as well as actual ex-post levels of life satisfaction and wage increases. 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 observed sample year being 2020. Third, we drop immigrants 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. The conditions they face are generally less applicable to our model, which requires access to the labor market and the ability to freely return to the country of origin.[28] Fourth, and most importantly, we need to ensure that predictions about immigrants' duration of stay are measured *at arrival*. Hence we restrict our sample to migrants who were first interviewed at most two years after their arrival in Germany.[29] This condition imposes the strongest restriction on sample sizes. Indeed, immigrants in the SOEP spend on average almost 9 years in Germany before they appear for the first time in the survey. Excluding refugees and ethnic Germans increases this period to almost 12 years. Finally, 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.

[28] Indeed, Table A2 shows that 95% of the sample of refugees and ethnic Germans intend to stay permanently. This proportion is of 63% in our main sample, which is composed of all other types of migrants (see Table 1).

[29] An implication of this is that for the first waves of the panel, migrants had to arrive in Germany in 1982 at the earliest.

Key variables We describe next how we construct each of the three central predictions made by migrants at arrival in our theoretical framework, i.e. return intentions, predictions about life satisfaction, and about wages. To capture whether these predictions were correct or not, we then compare them to expost outcomes (actual location, life satisfaction and wages), which results in measures of unexpected staying, life satisfaction and pessimism about wages.

A migrant is considered an unexpected stayer if they (i) initially have temporary intentions and (ii) are still in Germany in the long term.[30] Our main measure of unexpected staying, "UnexpS" is a binary variable which takes a value of 1 if the migrant is still in Germany in 2020 although they had expected a temporary stay. The variable takes a value of 0 if they either (i) left Germany as predicted, (ii) are still in Germany as predicted, or (iii) left Germany unexpectedly. We refer to these individuals as "other migrants", and show in Appendix B.4 that results are robust for two alternative definitions of unexpected staying. In the first alternative definition, we differentiate the "0" outcome between migrants with ex post correct predictions from migrants who left unexpectedly. In the alternative, we adjust our baseline definition using information on the difference of intended and realized years of stay for the subset of temporary migrants who provide this information. In this alternative definition, a migrant is thus an unexpected stayer if they are still in Germany and if their actual duration of stay goes beyond their intended duration.[31]

We proxy pessimism about utility by comparing information about migrants' predictions of future life satisfaction to their actual life satisfaction ex post.[32] The migrants in our sample are observed at several points in time, so that we have information about (i) their prediction at arrival in Germany about the life satisfaction they think they will have five years later, and (ii) the actual life satisfaction that they state five years later. Both variables are measured on a scale from 0 to 10. We construct our variable of $LS\_pessimism$ based on the difference between their actual life satisfaction ex post and the life satisfaction that they had predicted at arrival. The $LS\_pessimism$ score thus theoretically ranges between -10 and 10. For instance, a value of 10 corresponds to extreme pessimism, with a predicted life satisfaction of 0 at arrival, but an actual life satisfaction

[30] The SOEP 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) cannot 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).

[31] Our baseline definition overestimates the number of unexpected stayers if the latter includes immigrants who have temporary intentions but may not yet have reached their planned duration. In Appendix Table A3, we show that more than 82% of the immigrants with temporary intentions who provided an expected duration in years have exceeded their initial expectation which should mitigate this concern. Given that only a subsample of immigrants with temporary intentions also provide an expected duration in years, a definition conditional on the availability of expected years would significantly reduce our sample size.

[32] Measures of subjective well-being are known to correlate with migration decisions. For instance, Grimes and Wesselbaum (2021) show that differences between host and origin countries correlate with migration flows. Our contribution here pertains to the impact of (mis)predictions about life satisfaction on return migration.

of 10 five years later.

The measure of pessimism about wages follows a similar logic, by comparing the perceived likelihood of obtaining a wage raise in the following two years to the actual wage raise two years later.[33] The response modalities on beliefs about a potential wage raise changed in 1999, from a qualitative measure to a probabilistic measure. To reconcile this different measurement, we create a binary indicator which takes a value of 1 if the respondent expected a wage raise to occur either with the qualitative modalities "probably" or "definitely", or with a probability above 50%. Otherwise, the indicator takes a value of 0. 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 wage raise, whereas they did experience it ex post. Instead, a migrant turns out to be optimistic ($wage\_pessimism = -1$) if they expected higher wages, but did not obtain it ex post.

Finally, note that the construction of our pessimism measures requires that migrants in our sample have stayed in Germany (and in the SOEP) at least two and five years, respectively, in order to observe the realized wage evolution and life satisfaction. This potentially increases the probability that migrants become unexpected stayers in our sample. To rule out concerns of selection towards unexpected stayers in our baseline sample, we consider an alternative "unconstrained" sample in which we do not measure pessimism and only impose the observation of intentions at arrival and actual location in 2020. Table A4 in the Appendix allows to compare the proportions of unexpected stayers between these two samples. Unexpected staying is even more prevalent in the unconstrained sample (38.1%) than in our baseline sample (33.8%). It is thus unlikely that our baseline sample suffers from this type of selection.

Assumptions required to bridge theory and empirics Despite containing rich information, SOEP data still has important limitations that impact our analysis. In this Subsection, we summarize these limitations and explain the assumptions underlying our empirical analysis.[34]

First, the notion of unexpected staying in our model is based on a comparison between predicted and actual probabilities over a set of possible events, including the duration of stay and wages. In the SOEP data, predictions about a permanent stay are however binary ("yes, I will stay permanently" versus "no, I will stay temporarily"). The same applies to the expectation regarding a wage raise, whereas life satisfaction is measured on a discrete scale. Ideally, we would use probabilistic

[33] Predictions about a wage raise pertain to a shorter time horizon (2 years) than predictions about future life satisfaction (5 years). If wage predictions and actual wages had fewer missing values over time, we would have provided measures of wage pessimism over more than 2 years. Unfortunately, this is not feasible as it would further drastically reduce our sample sizes.

[34] In Appendix B.3, we provide additional information about the requirements that a hypothetical dataset should have met to ideally test our predictions.

predictions, i.e. statements by migrants about their perceived *probabilities* over a set of possible outcomes in the future (location, life satisfaction, wages). For example, instead of "Do you think that you will stay permanently in Germany? - Yes/No", a probabilistic prediction could have been obtained by asking "What is the probability that you will stay permanently in Germany? - ...%." Furthermore, we do not observe the probability of staying ex post, but instead its realization (or not). Although our definition of unexpected staying in the model is probabilistic, we only observe its realized counterpart in the data. While non-probabilistic predictions deprive us from useful variability within both explanatory factors (pessimism) and outcomes (unexpected staying), they do not introduce an estimation bias per se.

Second, the literature has established that migrants' return decisions may be affected by unforeseen life events.[35] We account for changes in marital status and in the number of children in our analysis, but have to assume that additional unobservable idiosyncratic shocks are balanced and independently distributed across migrants. More specifically, idiosyncratic shocks occurring to migrants during their stay should be uncorrelated with migrants' pessimism. Since the pessimism measure is created from comparing predictions at arrival with data several years in the future, it may capture part of the shocks and thus be correlated with outcomes.

Third, we must assume that unobserved macro factors also affect pessimistic and non-pessimistic migrants similarly. For instance, business cycles, changes in exchange rates, or political instability in the origin country, which occur during the migrants' stay in Germany, should not be correlated with predetermined pessimism about the destination country.

Finally, perhaps the biggest limitation of the SOEP data pertains to the fact that it does not allow us to fully capture the notion of *relative* pessimism about life satisfaction stated in the model.[36] Indeed, the prediction that we capture at arrival pertains to a time horizon of 5 years and is hence focused on life satisfaction in Germany.[37] Instead, the pessimism that we would like to measure should be *relative* between Germany and the migrant's origin country. A migrant exhibiting relative pessimism is indeed less optimistic about their likelihood to enjoy life in Germany *than in the origin country*.[38] Migrants may thus be optimistic about Germany but still exhibit relative pessimism if they are even more optimistic about the origin country than about Germany. While we have no way to measure relative pessimism because we have no predictions on the origin country, we show in Table 2 that migrants who are unexpected stayers are less optimistic about life in Germany than

[35] From a theoretical standpoint, we have argued in previous sections that such shocks cannot provide a systematic explanation to unexpected staying.

[36] Note that we do not face this issue with pessimism about wage prospects.

[37] For the subsample of migrants for which we observe the intended duration of stay (which is more demanding than the binary prediction temporary versus permanent), less than 20 percent of migrants state an intention to stay of less than five years. So there is a small minority of migrants who may consider this question through the lens of a return to the origin country.

[38] Indeed, we show under mispredicted long-term utility that unexpected staying results from $\widetilde{v}^d - \widetilde{v}^o \leq v^d - v^o$.

"other migrants".[39]

With these caveats in mind, we test the associations predicted by the model with the SOEP data in the next section.

7.2 Descriptive evidence

Table 2 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).[40] Cohorts of arrival in Germany are defined at the decade level: 1982-90; 1991-2000 and 2001-2010. We first present descriptive statistics of the whole sample in (1) and then split it between (2) unexpected stayers and (3) "other migrants" who either have left Germany or have correctly predicted to remain in Germany until at least 2020. Column (4) provides t-tests on the mean differences between subsamples (2) and (3).

Column (1) shows that 50% of all migrants in our sample had temporary intentions at arrival, whereas 78% are still in Germany in 2020. Out of the 204 migrants, 69 are unexpected stayers (i.e. they are in Germany in 2020 although they intended to leave the country at the beginning of their stay). Among the 135 migrants who are not unexpected stayers, 22% had planned to be temporary and have indeed left, 67% 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 by 2020). In terms of observable characteristics at arrival, unexpected stayers appear to be ex ante similar to other migrants.

Unexpected stayers remain in Germany on average 15.7 years more than they initially predicted, whereas other migrants stay only 2.2 years more than predicted. Both "unexpected stayers" and "other migrants" are optimistic about life in Germany. Optimism by migrants about the host country at arrival is sensible given the recent decision to emigrate there, and it is in line with previous research (Taylor et al., 2006). It can also be explained by unanticipated declines in actual life satisfaction that occur in the years that follow the arrival.[41] As previously mentioned, this does not imply that migrants are not *relatively* pessimistic compared to the origin country (see Section

[39] To claim that unexpected stayers exhibit higher *relative* pessimism than "other migrants", we would need to assume that the lower optimism about Germany by unexpected stayers (compared to other migrants) is not compensated by an even lower optimism about the origin country (compared to other migrants).

[40] The number of years of education is missing for 15% of our sample. In order to avoid losing these observations, we generate an indicator variable which is equal to 1 when the information on education is missing, and 0 otherwise. For these observations, we then replace the missing value for their education years by a value of 0, as the effect will be captured by the dummy. Note that using a different value for this imputation only changes the coefficient for the indicator variable in our regressions but leaves the other coefficients unaffected.

[41] Such declines might be explained by unmet aspirations and related concerns, for instance because migrants realize after some time that they earn lower wages than natives (Nikolova, 2015; Paparusso, 2021).

Table 2: Descriptive statistics

	All		Unexp.Stay= 1		Unexp.Stay= 0		Diff.	
	(1)		(2)		(3)		(4)	
Temp. intentions	0.50	(0.50)	1.00	(0.00)	0.22	(0.42)	0.78***	(20.44)
Still in Germany	0.78	(0.42)	1.00	(0.00)	0.67	(0.47)	0.33***	(8.19)
Years overstayed[1]	11.52	(9.96)	15.70	(9.07)	2.22	(3.37)	13.48***	(8.22)
LS-pessimism	-1.01	(1.97)	-0.56	(1.86)	-1.24	(1.99)	0.69*	(2.16)
W-pessimism	-0.16	(0.62)	0.00	(0.60)	-0.29	(0.61)	0.29*	(2.28)
Age	29.03	(10.03)	28.59	(10.53)	29.25	(9.80)	-0.66	(-0.43)
Female	0.52	(0.50)	0.45	(0.50)	0.56	(0.50)	-0.11	(-1.54)
Married	0.81	(0.39)	0.80	(0.41)	0.81	(0.39)	-0.02	(-0.30)
Education (in years)	8.50	(4.00)	8.41	(3.85)	8.54	(4.09)	-0.13	(-0.22)
Missing education[2]	0.15	(0.36)	0.14	(0.35)	0.16	(0.36)	-0.01	(-0.20)
Children	0.64	(0.96)	0.52	(0.76)	0.70	(1.05)	-0.18	(-1.41)
Chg. married	0.01	(0.41)	0.09	(0.37)	-0.02	(0.43)	0.11	(1.87)
Chg. children	1.03	(1.17)	1.19	(1.20)	0.96	(1.15)	0.23	(1.33)
82-90 cohort	0.23	(0.42)	0.28	(0.45)	0.21	(0.41)	0.07	(1.05)
91-00 cohort	0.60	(0.49)	0.54	(0.50)	0.64	(0.48)	-0.10	(-1.37)
Newspaper	2.10	(1.29)	1.96	(1.19)	2.17	(1.36)	-0.22	(-0.67)
Saving share	0.05	(0.09)	0.06	(0.09)	0.05	(0.09)	0.01	(0.72)
Observations	204		69		135		204	

Notes: [1]By design, the number of years of unexpected staying is only available for immigrants whose initial intention was to stay temporarily. Among them, the unexpected stayers have on average stayed 15.7 years longer than initially planned, while those who left Germany stayed 2.2 years longer than planned. Also note that the value of 15.7 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. [2]This variable equals 1 for individuals for whom information on number of education years is missing (15% of the sample) and 0 otherwise. For the latter, we assume their education years are equal to 0. Arrival cohorts 1982-90 and 91-2000 provide the share of respondents who arrived within a specific decade. The remaining 17% of respondents arrived between 2001 and 2010. ***,**,* denote significance at the 1, 5 and 10% level, respectively.

7.1). What is important to note here is that unexpected stayers are less optimistic about their life satisfaction than other migrants, with a t-statistic of the mean difference between both groups equal to 2.16.

When it comes to pessimism about wage prospects, unexpected stayers appear to be neutral, whereas "other migrants" tend to be optimistic as they overestimate their likelihood of obtaining a wage increase, with a t-statistic of the mean difference between both groups equal to 2.28. To sum up, we observe that unexpected stayers are the least optimistic at arrival about life satisfaction and wages in Germany.

Finally, Table 2 provides information related to integration investments i and savings s. Integration is proxied by the type of newspapers read by migrants on a scale from 1 (newspapers from the country of origin only) to 5 (German newspapers only). The saving rate is the proportion of savings in the household's monthly income. These variables suggest that unexpected stayers integrate slightly less (read less German newspapers) and save more than other migrants in the short term, though these differences are not statistically significant.

Unexpected staying and pessimism We estimate by OLS the following equation:

$$UnexpS_{i,2020} = c + \beta \cdot Pessimism_{i,t_i^0} + \gamma \cdot X_{i,t_i^0} + \delta \cdot Cohort\left(t_i^0\right) + \eta \cdot Shocks_i + \epsilon_i,$$

where $\text{UnexpS}_{i,2020}$ is a binary variable which equals 1 if migrant i (i) is still in Germany in 2020, and (ii) had predicted a temporary stay at their arrival in t_i^0. More specifically, $t_0^i \in [1982; 2010]$ varies across migrants and corresponds to the first year in which we observe migrant i's prediction about their stay. Since this prediction should be made as early as possible in the migration spell, we restrict our sample to migrants who stated it at the latest two years after their arrival in Germany.[42] $Pessimism_{i,t_i^0}$ is the measure of (wage or life satisfaction) pessimism in t_i^0. X_{i,t_i^0} is a set of migrant characteristics measured at arrival, $Cohort\left(t_i^0\right)$ is a dummy variable capturing the decade migrants arrived in Germany, and $Shocks_i$ is a set of life events which occurred between t_i^0 and 2020. Note that we study the effect of each of the two pessimism measures in two separate estimations.[43] The main parameter of interest β is expected to be positive as the model predicts that pessimism at arrival and probability of unexpected staying are positively correlated.

Tables 3 and 4 provide estimates of the determinants of unexpected staying. Column (1) controls for a set of individual characteristics measured at arrival (age, gender, marital status, number of

[42] We chose a two-year interval after arrival so as to balance sample sizes and being able to observe intentions early enough in their migration spell.

[43] Since life satisfaction and wage pessimism have many missing values, we only observe both forms of pessimism for 67 migrants. We thus use two different subsamples for each regression so as to maximize the number of available observations for each pessimism measure. Note that estimating the effects of pessimism separately should not alter our estimates (compared to a joint estimation) since the two pessimism measures are virtually uncorrelated (0.01) in the 67-observation sample.

Table 3: Unexpected staying and pessimism about life satisfaction

	(1)		(2)		(3)		(4)	
	Unexpected staying							
	b	se	b	se	b	se	b	se
LS-pessimism	0.042**	(0.02)	0.041**	(0.02)	0.042**	(0.02)	0.040**	(0.02)
Age	0.002	(0.00)	0.002	(0.00)	0.003	(0.00)	0.004	(0.00)
Female	-0.130*	(0.07)	-0.130*	(0.07)	-0.109	(0.07)	-0.105	(0.07)
Married	-0.012	(0.10)	-0.016	(0.10)	-0.030	(0.10)	0.016	(0.13)
Children	-0.025	(0.04)	-0.026	(0.04)	-0.049	(0.04)	-0.036	(0.04)
Education			-0.007	(0.02)	-0.001	(0.02)	0.004	(0.02)
Missing education			-0.065	(0.24)	-0.008	(0.24)	0.043	(0.24)
82-90 cohort					0.231	(0.14)	0.218	(0.14)
91-00 cohort					-0.041	(0.10)	-0.031	(0.10)
Chg. married							0.091	(0.12)
Chg. children							0.022	(0.04)
Constant	0.421***	(0.15)	0.492*	(0.27)	0.435	(0.29)	0.268	(0.33)
Observations	166		166		166		166	

Notes: Unexpected stayers (with UnexpS=1) are migrants who report an initial intention to leave but who are still in Germany in 2020. ***,**,* denote significance at the 1, 5 and 10% level, respectively.

children). Column (2) adds education (in years) and includes an indicator variable which takes value 1 when this information is missing (and 0 otherwise). Column (3) includes in addition cohort-of-arrival fixed effects (defined at the decade level: 1982-1990; 1991-2000 and 2001-2010 as reference). Column (4) adds controls for individual-level shocks that may affect return plans: changes in marital status and in the number of children.[44] Coefficient estimates are very stable throughout the different specifications. In both Table 3 and Table 4, the coefficient of pessimism is positive and significant, which highlights a positive association between pessimism (about both life satisfaction and wages) and unexpected staying. More specifically, the full specification in Table 3 (column 4) suggests that a migrant who had one unit of life-satisfaction pessimism (i.e. who had predicted at arrival a future life satisfaction one point below their actual life satisfaction 5 years later) is on average 4 percentage points more likely to stay unexpectedly in 2020. One standard deviation of pessimism

[44]For instance, a migrant who was married at arrival but who is single, divorced or widowed in the last survey year (i.e. not married) would thus have a change in marital status equal to -1. The change in number of children in the migrant's household is calculated as the difference between the maximum number of children ever observed in the household and the initial number of children at arrival. For example, 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 SOEP without children, the recorded change in the number of children is 3.

Table 4: Unexpected staying and wage pessimism

	(1)		(2)		(3)		(4)	
	Unexpected staying							
	b	se	b	se	b	se	b	se
W-pessimism	0.172**	(0.08)	0.172**	(0.08)	0.180**	(0.08)	0.184**	(0.08)
Age	-0.008	(0.01)	-0.009	(0.01)	-0.014*	(0.01)	-0.011	(0.01)
Female	-0.039	(0.11)	-0.049	(0.11)	-0.032	(0.11)	-0.032	(0.11)
Married	0.028	(0.11)	0.042	(0.11)	0.115	(0.12)	0.182	(0.14)
Children	-0.071	(0.06)	-0.072	(0.06)	-0.039	(0.07)	-0.044	(0.07)
Education			0.017	(0.03)	0.014	(0.03)	0.012	(0.03)
Missing education			0.142	(0.32)	0.155	(0.31)	0.140	(0.31)
82-90 cohort					-0.230	(0.16)	-0.212	(0.17)
91-00 cohort					-0.351**	(0.15)	-0.305*	(0.15)
Chg. married							0.161	(0.14)
Chg. children							0.006	(0.06)
Constant	0.698***	(0.20)	0.556*	(0.31)	0.894**	(0.35)	0.736*	(0.39)
Observations	105		105		105		105	

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 2020. ***,**,* denote significance at the 1, 5 and 10% level, respectively.

about life satisfaction (1.97 units of $LS - Pessimism$) is thus associated with a 7.9 percentage points higher probability of unexpected staying. Column 4 of Table 4 shows that a migrant who had failed to predict their significant wage increase tends to be 18.4 percentage points more likely to stay unexpectedly.

Appendix B.4 disaggregates the unexpected staying outcome by distinguishing two components: (1) a binary variable equal to 1 if the migrant had temporary intentions at arrival and (2) a binary variable equal to 1 if the migrant is still in Germany in 2020. Our main outcome variable is thus the product of these two binary variables. Migrants who were more 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). 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 4.

In Appendix B.5, 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. The link between pessimism and migrants' monthly saving share is negligible in both magnitude and significance.

Summing up, our empirical investigation is consistent with our theoretical results, as pessimism about utility and future wages is positively and significantly associated with unexpected staying. We acknowledge however that this analysis would benefit from richer data. In particular, direct tests of life satisfaction pessimism would require country-specific life satisfaction predictions. Indeed, our measure of pessimism is one of greater pessimism about the destination than the origin country – thus to measure relative pessimism, one should observe migrants' predictions about utility in both countries separately.

8 Conclusion

Empirical evidence shows that migrants tend to systematically underestimate their propensity to stay in the long term. With a simple theoretical framework, we show that this unexpected staying results from pessimistic misperceptions about the host country relative to the origin country at arrival. We identify biases that systematically give rise to such pessimism, such as minsinference, 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 the probability of staying unexpectedly in the long term. Specifically, ex-ante pessimism about life satisfaction and about future wages, based on the comparison between predictions and realizations ex post, leads to a higher likelihood of staying unexpectedly.

While populations and governments of host countries often fear that immigrants may be too optimistic about the host country, we find that it is pessimism about the host country relative to the origin country which matters for unexpected staying and low integration. Optimism about the host country in the short term provides in fact incentives to integrate, whereas an excess of optimism about the origin country in the long term (i.e. migrants idealizing their life in case of return) deters integration and fosters unexpected staying. This suggests that migrants arriving in the host country would benefit from clearer signals about their long-term prospects in both locations, i.e. in case of stay as well as in case of return. This is in line with a recent a growing strand of literature highlighting the importance of information and potential misperceptions on prospective migrants' decisions (Bah and Batista, 2020; Batista and McKenzie, 2021; Tjaden and Gninafon, 2022; Battiston et al., 2022). These papers find that incomplete information and risks related to migrating, to unemployment and to the possibility to stay in the host country can affect migration decisions, and that information treatments can help their decision-making.

Misperceptions are a central feature of this paper, and our approach involves certain modeling choices. First, we show that misperceptions can be explained by certain behavioral biases. Alternative mechanisms, such as institutional or social factors (biased reporting, information campaigns or aggregation in networks) might also generate the misperceptions studied in this paper. Second, while our framework can explain the origins of misperceptions, we do not aim to explain their evolution throughout the migration spell. Third, an important distinction between misperceptions and uncertainty should be made. Our model focuses on one dimension of uncertainty (long-term wages) but other sources of uncertainty, for instance about life events (such as getting married or divorced, or having children), may also affect the integration process and return intentions. Yet, if perceptions about the distribution of these shocks are correct ex ante, uncertainty about these shocks cannot systematically generate unexpected staying. Indeed, only a structural gap between beliefs and realizations can explain its systematic nature, so uncertainty per se is not essential to explain unexpected staying. We do not model uncertainty about the integration process, but allow migrants to have misperceptions about it through projection bias and narrow bracketing.

In the empirical analysis, plausibly exogenous variation of pessimism would have allowed us to estimate its causal effect. 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 might serve to identify it. Such designs would however face the challenging constraint that unexpected staying requires a long-term data collection process. More generally, better data that directly elicits probabilistic beliefs and predictions of migrants would help to measure misperceptions, and provide better evidence for the actual degree of unexpected staying.

References

Abramitzky, R., Boustan, L., and Eriksson, K. (2019). To the new world and back again: Return migrants in the age of mass migration. *ILR Review*, 72(2):300–322.

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

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

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. *International Migration*, 43(3):131–154.

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. (2019). An Experiment on Time Preference and Misprediction in Unpleasant Tasks. *Review of Economic Studies*, 86(3):941–975.

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.

Bah, T. L. and Batista, C. (2020). *Why Do People Migrate Irregularly?: Evidence from a Lab-in-the-field Experiment in West Africa.* Helen Kellogg Institute for International Studies.

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

Batista, C. and McKenzie, D. J. (2021). Testing classic theories of migration in the lab. Technical Report No.16469.

Battiston, G., Corno, L., and Ferrara, E. L. (2022). Informing risky migrants: Evidence from a field experiment in guinea. *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.

Bolognani, M. (2007). The myth of return: Dismissal, survival or revival? A Bradford example of transnationalism as a political instrument. *Journal of Ethnic and Migration Studies*, 33(1):59–76.

Borjas, G. J. and Bratsberg, B. (1996). Who leaves? The outmigration of the foreign-born. *Review of Economics and Statistics*, 78(1):165–176.

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. mimeo.

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.

Delavande, A. (2014). Probabilistic expectations in developing countries. *Annual Review of Economics*, 6(1):1–20.

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.

Esponda, I. and Vespa, E. (2014). Hypothetical thinking and information extraction in the laboratory. *American Economic Journal: Microeconomics*, 6(4):180–202.

Exley, C. L. and Kessler, J. B. (2018). Equity Concerns are Narrowly Framed. NBER Working Papers 25326, National Bureau of Economic Research, Inc.

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.

Fasani, F., Frattini, T., and Minale, L. (2021). Lift the Ban? Initial Employment Restrictions and Refugee Labour Market Outcomes. *Journal of the European Economic Association*, 19(5):2803–2854.

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.

Goebel, J., Grabka, M. M., Liebig, S., Kroh, M., Richter, D., Schröder, C., and Schupp, J. (2019). The german socio-economic panel study (soep). *Jahrbücher fÃijr Nationalökonomie und Statistik / Journal of Economics and Statistics*, 239(2):345–360.

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.

Grimes, A. and Wesselbaum, D. (2021). The Role of Subjective Wellbeing in Cross-Border Migration. In Kourtit, K., Newbold, B., Nijkamp, P., and Partridge, M., editors, *The Economic Geography of Cross-Border Migration*, pages 217–243. Springer International Publishing.

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).

Ivlevs, A. (2015). Happy Moves? Assessing the Link between Life Satisfaction and Emigration Intentions. *Kyklos*, 68(3):335–356.

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.

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.

Manski, C. F. (2004). Measuring expectations. *Econometrica*, 72(5):1329–1376.

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. *Annual Review of Economics*, 4(1):511–540.

Nikolova, M. (2015). Migrant well-being after leaving transition economies. *The IZA World of Labor*, pages 195–195.

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.

Paparusso, A. (2021). *Immigrant Integration in Europe. Human Well-Being Research and Policy Making.* Springer.

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. *The World Bank Economic Review*, 34(1):1âĂŞ27.

Sinatti, G. (2011). Mobile transmigrants or unsettled returnees? myth of return and permanent resettlement among senegalese migrants. *Population, Space and Place*, 17(2):153–166.

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.

Taylor, P., Funk, C., and Craighill, P. (2006). Once Again, the Future AinâĂŹt What It Used to Be. Technical report, Washington: Pew Research Center.

Tjaden, J. and Gninafon, H. (2022). Raising Awareness About the Risk of Irregular Migration: Quasi-Experimental Evidence from Guinea. *Population and Development Review*, forthcoming.

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.

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 the 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 (s,i), 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 of Lemma 1

First note that $w_R(s,i)$ is defined by the implicit function $v^d(s,i;w_R) - v^o(s) = 0$, or equivalently $u^d\left(c^d(s), I^d(i)\right) - u^o\left(c^o(s), I^o\right) = 0$, where $c^d = w_R + s$, $I^d = I_1 + i$, $c^o = w^o + xs$, and I^o is exogenous. Applying the implicit function theorem to this equality, we obtain 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 evaluated at values of $\left(c^d,I^d\right)$ which make the migrant indifferent between the two locations, that is, $u^d\left(c^d,I^d\right) = u^o\left(c^o,I^o\right)$. Since for all i, $I^o \geq I^d$, it must be that $c^d \geq c^o$ (otherwise the migrant would 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 fact that $I^o \geq I^d$ reinforces the result.

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

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

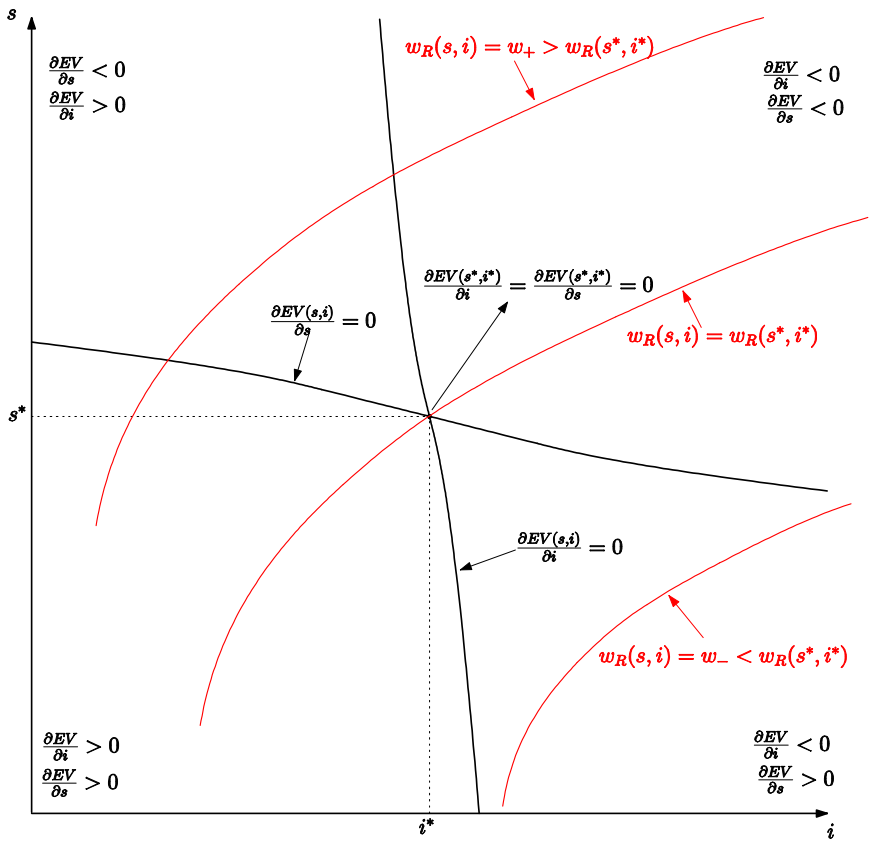

Figure 1: Representation of the reservation wage and first order conditions in the (s,i) space

■

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* $EV(s,i) = EV(s,i;v_1,v^d,v^o,F)$, *and* $\left(\widetilde{s},\widetilde{i}\right) = \arg\max_{s,i} \widetilde{EV}(s,i)$, *where* $\widetilde{EV}(s,i) = EV\left(s,i;\widetilde{v}_1,\widetilde{v}^d,\widetilde{v}^o,\widetilde{F}\right)$.[45] *Then* (s^*,i^*) *must satisfy*

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

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

[45] 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 of Lemma 2

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:

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

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

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

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 (s,i), (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.[46] 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.

The combination of Lemma 1 and 2 leads to the following preliminary result. This result, which allows for a wide range of simultaneous misperceptions, states that if a migrant (of any characteristics) is *generally pessimistic*, then they are *more likely to integrate less and to save more* than they would if they had correctly perceived their core parameters. By general pessimism, we mean that the migrant underestimates the lifetime 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. This is formally defined here as G-pessimism.

Definition 5. *A migrant exhibiting general pessimism (**G-Pessimism**) is expected to overestimate the benefits of savings and to underestimate the benefits of integration:* $\frac{\partial E\widetilde{V}(s,i)}{\partial i} < \frac{\partial EV(s,i)}{\partial i}$ *and* $\frac{\partial E\widetilde{V}(s,i)}{\partial s} > \frac{\partial EV(s,i)}{\partial s}$ *for all* (s,i).

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

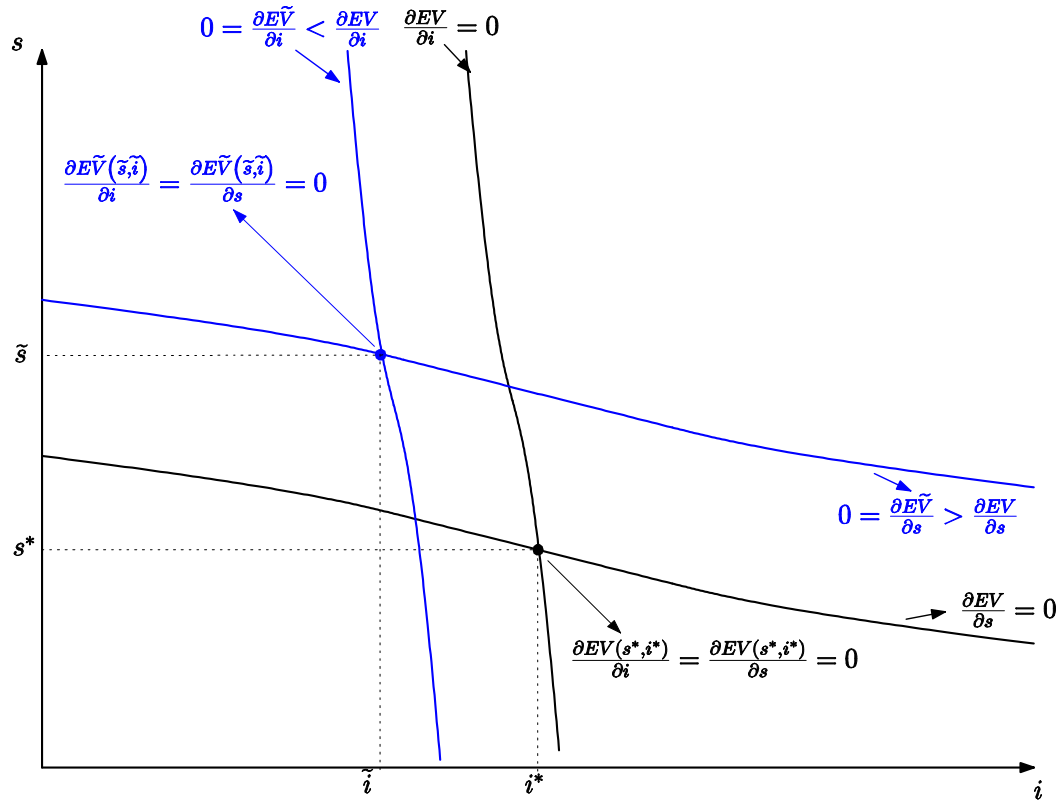

Figure 2: Representation of misperceived first order conditions and predicted choices in the (s,i) space

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). We can now state the first formal result.

Proposition 3. *Consider all possible migrants z who share the same set of misperceptions $\left(\widetilde{v}_1, \widetilde{v}^d, \widetilde{v}^o, \widetilde{F}\right) \neq \left(v_1, v^d, v^o, F\right)$. If these misperceptions reflect* ***G-Pessimism****, then for all z we have that $\widetilde{i} < i^*$ and $\widetilde{s} > s^*$.*

Proof of Proposition 3

To prove that G-pessimism is sufficient, we build on our graphical representation of the migrant's choice depicted in Figure 1, and we 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. We 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 $\frac{\partial EV}{\partial s} = 0$.

We take (s^*, i^*) as our reference point. Since pessimism means that $\frac{\partial E\widetilde{V}(s,i)}{\partial i} < \frac{\partial EV(s,i)}{\partial i}$ for all (s,i), and since (s^*, i^*) is such that $\frac{\partial EV(s^*,i^*)}{\partial i} = 0$, we conclude that $\frac{\partial E\widetilde{V}(s^*,i^*)}{\partial i} < 0$. Since $\frac{\partial^2 E\widetilde{V}}{\partial i^2} < 0$, 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 $\frac{\partial EV(s,i)}{\partial i} = 0$. We then apply the same reasoning with $\frac{\partial E\widetilde{V}(s,i)}{\partial s}$. Since $\frac{\partial E\widetilde{V}(s,i)}{\partial s} > \frac{\partial EV(s,i)}{\partial s}$ for all (s,i), and (s^*, i^*) is such that $\frac{\partial EV(s^*,i^*)}{\partial s} = 0$, we conclude that $\frac{\partial E\widetilde{V}(s^*,i^*)}{\partial s} > 0$. 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^*).[47] Since this intersection is by definition $\left(\widetilde{s}, \widetilde{i}\right)$, we can conclude that $\widetilde{s} > s^*$ and $\widetilde{i} < i^*$.

■

A.2 Mispredicted probabilities

Proof of Proposition 1

The formal statement which needs to be proven is: $\widetilde{F}(w^d) > F(w^d)$ for all w^d if and only if $\widetilde{F}\left(w_R\left(\widetilde{s}, \widetilde{i}; v^d, v^o\right)\right) > F\left(w_R\left(\widetilde{s}, \widetilde{i}; v^d, v^o\right)\right)$ for all $\left(z, v_1, v^d, v^o, F\right)$, where $(\widetilde{s}, \widetilde{i})$ maximizes $EV\left(s, i; v_1, v^d, v^o, \widetilde{F}\right)$. Proving the sufficiency condition ($\Longrightarrow$) is trivial: under W-pessimism, $\widetilde{F}(w^d) > F(w^d)$ 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 $\widetilde{F}\left(w_R\left(\widetilde{s}, \widetilde{i}; v^d, v^o\right)\right) > F\left(w_R\left(\widetilde{s}, \widetilde{i}; v^d, v^o\right)\right)$ 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 $\widetilde{F}(w^d) > F(w^d)$ for all w^d, which is our definition of W-pessimism. ■

A.3 Proposition 4

Proposition 4. *Consider all possible migrants $\left(z, v_1, v^d, v^o\right)$ with the same misperceptions $\widetilde{F}(\cdot)$. If $\widetilde{F}(\cdot)$ satisfies W-pessimism, then for all $\left(z, v_1, v^d, v^o\right)$ we have that $\widetilde{s} > s^*$ and $\widetilde{i} < i^*$.*

Proof of Proposition 4

Proposition 4 is a direct application of Proposition 3 once we show that W-pessimism implies $\frac{\partial E\widetilde{V}}{\partial i} < \frac{\partial EV}{\partial i}$ and $\frac{\partial E\widetilde{V}}{\partial s} > \frac{\partial EV}{\partial s}$.

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

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

[47] 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.

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

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

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

The first term of this expression is strictly 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: $\int_{w_R}^{\infty} \frac{\partial v^d}{\partial s}\left(\widetilde{f}\left(w^d\right) - f\left(w^d\right)\right) dw^d$

$$= \left[\frac{\partial v^d}{\partial s}\left(\widetilde{F}\left(w^d\right) - F\left(w^d\right)\right)\right]_{w_R}^{\infty} - \int_{w_R}^{\infty} \frac{\partial^2 v^d}{\partial s \partial w^d}\left(\widetilde{F}\left(w^d\right) - F\left(w^d\right)\right) dw^d$$

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

Combining both terms, we obtain:

$$\frac{\partial E\widetilde{v}_2}{\partial s} - \frac{\partial Ev_2}{\partial s} = \left(\frac{\partial v^o(s)}{\partial s} - \frac{\partial v^d(s,i;w_R)}{\partial s}\right)\left(\widetilde{F}(w_R) - F(w_R)\right) + \int_{w_R}^{\infty}\left(-\frac{\partial^2 v^d}{\partial s^2}\right)\left(\widetilde{F}\left(w^d\right) - F\left(w^d\right)\right) dw^d.$$

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.

A.4 Extension: integration improves the distribution of wages

Our baseline model focuses on migrants' utility of being integrated (based on a function u(c; I)). We now consider an extension of the model in which migrants' integration can improve their wage

distribution: $F\left(w^d; I^d\right)$, 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:

$$\frac{\partial EV}{\partial i} = \frac{\partial v_1(s,i)}{\partial i} + \delta \int_{w_R}^{\infty} \frac{\partial v^d(s,i)}{\partial i} f\left(w^d\right) dw^d + \int_{w_R}^{\infty} v^d(s,i) \left(\frac{\partial f\left(w^d\right); I^d}{\partial i}\right) dw^d.$$ [48]

We focus on the most 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 step **(1)** of the proof of **4** in Section A.3 ($\frac{\partial E\widetilde{v}_2}{\partial i} - \frac{\partial Ev_2}{\partial i}$) contains an additional term, namely $\int_{w_R}^{\infty} v^d(s,i) \left(\frac{\partial\left(\widetilde{f}\left(w^d;I^d\right) - f\left(w^d\right);I^d\right)}{\partial i}\right) dw^d$. Underestimating the benefits of integration on wages makes $\frac{\partial E\widetilde{v}_2}{\partial i} - \frac{\partial Ev_2}{\partial i}$ even more negative, since the additional term is also negative, as we can show by integrating by parts:

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

The first term on the right hand side is negative since

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

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 the effect of integration on wage prospects is their only misperception, the reasoning applies identically, which also results in $\widetilde{i} < i^*$ and $\widetilde{s} > s^*$.

■

A.5 Mispredicted long-term utility

Proof of Proposition 2

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

We 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, we first show that LT-pessimism implies $\widetilde{w}_R(s,i) > w_R(s,i)$ for any (s,i). We apply LT-pessimism for $w^d = w_R$: $\widetilde{v}^d(s,i;w_R(s,i)) - \widetilde{v}^o(s) < v^d(s,i;w_R) - v^o(s) = 0$. 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), $\widetilde{v}^d(s,i;w_R(s,i)) - \widetilde{v}^o(s) < \widetilde{v}^d(s,i;\widetilde{w}_R(s,i)) - \widetilde{v}^o(s)$, or equivalently $\widetilde{v}^d(s,i;w_R(s,i)) < \widetilde{v}^d(s,i;\widetilde{w}_R(s,i))$. This inequality is true if and only if $w_R(s,i) < \widetilde{w}_R(s,i)$.

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

We 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 $F\left(w_R\left(\widetilde{s},\widetilde{\imath};\widetilde{v}^d,\widetilde{v}^o\right)\right) > F\left(w_R\left(\widetilde{s},\widetilde{\imath};v^d,v^o\right)\right)$ for all (v_1,F,z), then $\widetilde{v}^d\left(s,i;w^d\right) - \widetilde{v}^o(s) < v^d\left(s,i;w^d\right) - v^o(s)$ for all (s,i). First note that $F(\widetilde{w}_R(\widetilde{s},\widetilde{\imath})) > F(w_R(\widetilde{s},\widetilde{\imath}))$ implies $\widetilde{w}_R(\widetilde{s},\widetilde{\imath}) > w_R(\widetilde{s},\widetilde{\imath})$, and the latter inequality implies that $\widetilde{v}^d(s,i;\widetilde{w}_R(\widetilde{s},\widetilde{\imath})) > \widetilde{v}^d(s,i;w_R(\widetilde{s},\widetilde{\imath}))$. Thus, $\widetilde{w}_R(\widetilde{s},\widetilde{\imath}) > w_R(\widetilde{s},\widetilde{\imath})$ implies $\widetilde{v}^d(\widetilde{s},\widetilde{\imath};\widetilde{w}_R(\widetilde{s},\widetilde{\imath})) - \widetilde{v}^o(\widetilde{s}) > \widetilde{v}^d(\widetilde{s},\widetilde{\imath};w_R(\widetilde{s},\widetilde{\imath})) - \widetilde{v}^o(\widetilde{s})$, where the left hand side of the latter inequality is by definition equal to 0, and thus equal to $v^d(\widetilde{s},\widetilde{\imath};w_R(\widetilde{s},\widetilde{\imath})) - v^o(\widetilde{s})$. Substituting this LHS, we obtain $v^d(\widetilde{s},\widetilde{\imath};w_R(\widetilde{s},\widetilde{\imath})) - v^o(\widetilde{s}) > \widetilde{v}^d(\widetilde{s},\widetilde{\imath};w_R(\widetilde{s},\widetilde{\imath})) - \widetilde{v}^o(\widetilde{s})$. Since $(\widetilde{s},\widetilde{\imath})$ is defined for all possible (v_1,F,z), no restrictions are imposed on the possible values of $(\widetilde{s},\widetilde{\imath})$, so this condition must hold for any $(\widetilde{s},\widetilde{\imath})$. To conclude, simply note that this inequality for any $(\widetilde{s},\widetilde{\imath})$ is identical to the definition of LT-pessimism.■

A.5.1 Proposition 5

Before stating Proposition 5, we 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 both integration and savings in the destination country ($\partial\widetilde{v}^d/\partial s < \partial v^d/\partial s$ and $\partial\widetilde{v}^d/\partial i < \partial v^d/\partial i$ for all s and i).*

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

Proposition 5. *Consider all possible migrants (z,v_1,F) with the same misperceptions $\left(\widetilde{v}^d,\widetilde{v}^o\right)$. If $\left(\widetilde{v}^d,\widetilde{v}^o\right)$ satisfies LTR-pessimism, then for all (z,v_1,F) we have that $\widetilde{s} > s^*$ and $\widetilde{i} < i^*$.*

Proof of Proposition 5

To prove that LTR-pessimism implies $\widetilde{s} > s^*$ and $\widetilde{i} < i^*$, it suffices to show that LTR-pessimism, that is $\partial\widetilde{u}^d/\partial c < \partial u^d/\partial c$ and $\partial\widetilde{u}^d/\partial I < \partial u^d/\partial I$ for all c and I, implies that $\frac{\partial E\widetilde{V}(s,i)}{\partial i} < \frac{\partial EV(s,i)}{\partial i}$

and $\frac{\partial E\widetilde{V}(s,i)}{\partial s} > \frac{\partial EV(s,i)}{\partial s}$ for all (s,i). Using Proposition 3, this twofold inequality yields $\widetilde{s} > s^*$ and $\widetilde{i} < i^*$.■

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: $F\left(w_R\left(\widetilde{s},\widetilde{i};v^d,v^o\right)\right) > F\left(w_R\left(s^*,i^*;v^d,v^o\right)\right)$. However, when they stay they are obviously worse off.

A.6 Mispredicted short-term utility

In Subsection A.6.1, we show the conditions about the (utility) costs of short-term actions that lead to unexpected staying. In Subsection A.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 the conditions on misperceptions about short-term costs of integration and savings, and thus does not lead to unexpected staying for all migrant types.

A.6.1 Short-term Pessimism and Unexpected Staying

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

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

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

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

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

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

$$|\partial\widetilde{v}_1/\partial i| = \widetilde{k}'(i) > k'(i) = |\partial v_1/\partial i|,$$

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

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

Overestimating the cost of integration is easily seen as a form of pessimism about the destination country. The reason we may also 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* $\left(z, v^d, v^o, F\right)$ *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* $\widetilde{v}_1$ *satisfies ST-pessimism.*

Proof of Proposition 6

To prove the necessity of ST-pessimism 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 can lead to either over- or underestimations of short-term returns depending on migrant characteristics.

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

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

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

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

$\Longleftarrow$ We prove the following:

Claim 1. *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 migrant preferences is that* $\widetilde{k}'(i) > k'(i)$ *and* $\widetilde{u}_1'(c) < u_1'(c)$*.*

Suppose that we have a bias and the condition of the claim does not hold, so that there is some separable v_1 with either $\widetilde{k}'(i) < k'(i)$ for all i or $\widetilde{u}'(c) < u'(c)$ for all c.[49] Then it is enough to show that we can find a distribution of wages $F(\cdot)$ such that $\widetilde{w}_R < w_R$ and such that intervals around these indifference wages are in the support of F – i.e. these migrants are unexpected *leavers*.

We only consider migrant preferences that are separable: $v_1(s,i) = u(c_1(s)) + g(I) - k(i)$, where $g(\cdot)$ is the utility from integration, and $k(\cdot)$ is the cost of increasing integration. Moreover, 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, utility functions are

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

the same in the two locations. If we can show the condition is necessary for this restricted class of preferences, then the condition is certainly necessary for more general preferences.

Suppose the bias does not satisfy the condition for $v_1(s,i)$. Then consider the following family of preferences parameterized by μ:[50]

- $v_1(s,i;\mu) \equiv \mu \cdot u(c(s)) + g(I) - k(i)$
- $v^o(s;\mu) \equiv \mu \cdot u(c^o(s)) + g(I^o)$
- $v^d(s,i,w;\mu) \equiv \mu \cdot u(c^d(s,w)) + g(I^d(i))$

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. Note that, as explained previously, actual decisions under mispredicted short-term preferences are ex post optimal. In order to avoid confusions, in this proof we introduce the notations $\hat{s}, \hat{\imath}$ which represent actual decisions made in period 1 in various hypothetical (constrained) scenarios.

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 $\widetilde{w}_R^N(\widetilde{s},\widetilde{\imath},\mu,F) < w_R^N(\hat{s},\hat{\imath},\mu,F)$, where the superscript N highlights that there is "No return" – that this w_R is a the indifference wage for a different setup, where the person is not allowed to return. It captures the indifference wage where a migrant who thought they would never be able to return is surprised in period 2 with the choice that they can return. 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 – that is, there is no wage for which they would want to return – with this wage distribution, their indifference wage is the same as when they face an exogenous constraint of not returning, as this constraint is not binding. Hence by the first step, we can find a $\mu_0 = \mu(F_0)$ such that $\widetilde{w}_R(\widetilde{s}_0,\widetilde{\imath}_0,\mu_0,F_0) = \widetilde{w}_R^N(\widetilde{s}_0,\widetilde{\imath}_0,\mu_0,F_0) < w_R^N(\hat{s}_0,\hat{\imath}_0,\mu_0,F_0) = w_R(\hat{s}_0,\hat{\imath}_0,\mu_0,F_0)$. In this unconstrained case, there are no wages that can occur under which the migrant would actually return. Therefore we finally define F_ε as a wage distribution that puts total weight ε on wages below $\bar{w}$, distributed

[50] 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.

uniformly, and we show that as $\varepsilon \to 0$, $\widetilde{w}_R(\widetilde{s}(\varepsilon), \widetilde{i}(\varepsilon), \mu_0, F_\varepsilon) \to \widetilde{w}_R(\widetilde{s}_0, \widetilde{i}_0, \mu_0, F_0)$ and $w_R(F_\varepsilon)$ and similarly for $w_R(\varepsilon) \to w_R(0)$. This step completes 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.

We now define the different settings.

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 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 $\widetilde{s}^N$ are the anticipated savings under no return and $\hat{s}^N$ are the actual savings under no return. Then we have:

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

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

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 disallow return there is no indirect impact either through the actual return decision. Therefore we can rewrite this maximization over both (s, i) jointly as two independent maximizations over i and s:

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

and similarly for anticipated choices:

$$\begin{aligned}
\widetilde{s}^N(F) &= \arg\max_s \widetilde{u}(S_0 - s) + \delta \int_0^\infty u_d(s + w) f(w)dw \\
\widetilde{i}^N &= \arg\max_i \delta g(I + i) - \widetilde{k}(i)
\end{aligned}$$

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

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

Before diving in, we define a few values to highlight that they do not depend on our later choices of μ. If $\widetilde{k}'(i) \geq k'(i)$ for all i, then set $\mu_L = 1$. If not, then we show that since $\widetilde{k}'(i) < k'(i)$ for all

i (we assume that the bias distorts marginals in one direction), then we must have that $\widetilde{i}^N > \hat{i}^N$. Suppose not, so that we have $\widetilde{i}^N \leq \hat{i}^N$. Then (assuming FOCs yield the optimal choices) we have that

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

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

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 $(0, \frac{\Delta(\widetilde{i}^N, \hat{i}^N)}{u(x\cdot(\bar{s}+w_o))})$, which is a non-empty interval of strictly positive numbers since $\hat{i}^N < \widetilde{i}^N$.

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. While this can be shown formally, intuitively it says that, no matter how little integrated the person is and how low (or high) their savings are, there is a wage high enough to ensure that the migrant always stays in the destination country.

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 consumption, which favors the destination country. By construction of F_ε, a migrant with μ_L-preferences stays for every wage $w > \bar{w}$:

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

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

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

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

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

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

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

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

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

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

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 $\hat{s}(F_0) = \hat{s}^N(F_0) > \widetilde{s}^N(F_0) = \widetilde{s}(F_0)$. Let us write $\widetilde{w}_R(\mu) \equiv w_R(\widetilde{s}(F_0), \widetilde{i}, F_0, \mu)$ 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 $\mu \geq \mu_L$:

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

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

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

where the last line uses similar limits for w_R as for $\widetilde{w}_R$. Since $\hat{s}^N > \widetilde{s}^N$, this limit is larger than 0. Thus we can pick a sufficiently large μ_0 s.t. $\widetilde{w}^N(\mu_0) < w^N(\mu_0)$. 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 $\widetilde{u}'(c) > u'(c)$ for all c, then we are done by applying step 1. Thus we can assume that $\widetilde{u}'(c) \leq u'(c)$ for all c, so that $s^N \leq \widetilde{s}^N$ for all wage distributions by the same argument as in step 1, but with the role of $\widetilde{u}$ and u reversed. Moreover, as we proved in the preliminaries, we have that $\widetilde{k}'(i) < k'(i)$ which implies that $\hat{i}^N < \widetilde{i}^N$.

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

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

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

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

where the last line follows since $\widetilde{s}^N - \hat{s}^N \geq 0$. 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.■

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 $z = (s_0, I_1, x, I^o, w^o)$, 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 $(\widetilde{s}, \widetilde{\imath})$ to their actual actions, which are ex post the optimal actions (s^*, i^*).

Proposition 7. *Consider all possible migrants $\left(z, v^d, v^o, F\right)$ with the same misperception $\widetilde{v}_1$. If $\widetilde{v}_1$ satisfies ST-pessimism, then for all $\left(z, v^d, v^o, F\right)$ we have that $\widetilde{s} > s^*$ and $\widetilde{\imath} < i^*$.*

Proof of Proposition 7

In the absence of other misperceptions, ST-pessimism is equivalent to $\frac{\partial E\widetilde{V}(s,i)}{\partial i} < \frac{\partial EV(s,i)}{\partial i}$ and $\frac{\partial E\widetilde{V}(s,i)}{\partial s} > \frac{\partial EV(s,i)}{\partial s}$ for all (s, i), hence by applying Proposition 3, we can directly show that ST-pessimism is sufficient for $\widetilde{s} > s^*$ and $\widetilde{\imath} < i^*$.■

A.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:

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

where $g_1(s, i)$ is the period-1 utility function. Equivalently, we can rewrite this programme by dividing the objective function by β to obtain

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

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

$$\widetilde{v}_1(s, i) = \frac{\beta}{\widetilde{\beta}} v_1(s, i).^{51}$$

[51] 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 $\widetilde{v}_1(s, i) = \frac{1}{\widetilde{\beta}} z_1(s, i) = \frac{\beta}{\widetilde{\beta}} v_1(s, i)$.

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 exist some migrant types for which present bias implies unexpected *leaving* rather than unexpected staying.

B Appendix: Data

B.1 Sample selection criteria and effect on sample sizes

Table A1: Sample selection criteria and sample sizes

	LS-pessimism (1)	W-pessimism (2)
Immigrant	29793	29793
+ Pessimism & Return int.	5332	6914
+ interviewed pre-2010	3330	3604
+ no refugee nor ethnic Germ.	2327	2771
+ years since arrival ≤2	178	109
+ age at arrival ≥18	166	105

Note: This table presents the evolution of sample sizes as selection criteria are sequentially applied. Column (1) refers to the sizes of samples containing information on LS-pessimism, and column (2) refers to the W-pessimism samples.

Table A1 provides the impact of different selection criteria on our sample sizes. Column (1) refers to the construction of the LS-pessimism sample whereas column (2) refers to the W-pessimism sample. The SOEP includes 29.793 immigrants that reply at least once to the survey. We then condition on the availability of the pessimism measure and return intention. Next, we exclude ethnic Germans and refugees, given the different institutional setting that they face. Then, we limit the sample to individuals interviewed within two years after their arrival in Germany, given that we want to avoid having individuals in the sample who made integration decisions and acquired sufficient information about their career prospect. Finally, we exclude individuals that were under the age of 18 at arrival.

Table 1 excluded ethnic Germans and refugees, since to be aligned with our theory, our population of interest should be able to freely return to their home country. The relevance of this condition is highlighted by Table A2, which focuses exclusively on refugees and ethnic Germans. It shows that 94,9% of these migrants intended to stay permanently. Among the few who intended to

Table A2: Return intentions and actual location in 2020 of refugees and ethnic Germans

	Actual location in 2020		
Initial intentions	in Germany	left Germany	
Permanent stay	12,539	304	12,843
Temporary stay	644	33	677
	13,183	337	13,520

Note: SOEP data for immigrants who are either refugees or ethnic Germans, surveyed between 1984 and 2020. Reported intentions to stay (temporarily or permanently) are collected from the first reply to the related survey question.

stay temporarily, 95,1% are unexpected stayers, which can be attributed, on top of the mechanisms studied in this paper, to refugees' mere impossibility to return.

Table A3: Return intentions and location in 2020:
subsample with available information on intended years of stay

	All Migrants	Excl. Refugees & Ethnic Germans
	(1)	(2)
Temporary intentions	3.833	3.624
Exceeded intended duration	3.121	2.972

This table shows the number of immigrants with temporary intentions who provide information on their number of intended years of stay, and among those, the number of migrants who already exceeded their intended duration. Column 1 concerns all migrants including refugees, column 2 excludes refugees and ethnic Germans.

In Table A3, we focus on the subsample of immigrants with temporary intentions who provide information on their intended duration of stay (in years). Column 2 (all migrants except for refugees and ethnic Germans) shows that 82% already stayed longer than their intended duration. This mitigates the concern that the phenomenon of unexpected staying is overestimated with our baseline definition, since only 18% of stayers may return in time with respect to their initial plan.

B.2 Measures of pessimism do not alter sample composition

Table A4 provides descriptive statistics for the unconstrained sample, i.e. the sample of individuals that provide a return intention but not necessarily information to construct any of the two pessimism variables. The availability of a pessimism variable does not alter the sample composition, in particular in terms of unexpected staying: if anything, unexpected staying is more prevalent in the unconstrained (38.1%) than in the constrained sample (33.8%).

Table A4: Descriptive statistics - Sample meeting demographic conditions

	All		Unexp.Stay= 1		Unexp.Stay= 0		Diff.	
Temp. intentions	0.49	(0.50)	1.00	(0.00)	0.23	(0.42)	0.77***	(19.75)
Still in Germany	0.76	(0.43)	1.00	(0.00)	0.62	(0.49)	0.38***	(12.23)
Years overstayed	11.52	(9.96)	16.08	(8.87)	2.16	(3.29)	13.92***	(8.66)
Age	29.18	(11.32)	28.17	(11.20)	29.79	(11.37)	-1.62	(-1.38)
Female	0.60	(0.49)	0.59	(0.49)	0.61	(0.49)	-0.02	(-0.47)
Married	0.78	(0.41)	0.81	(0.39)	0.77	(0.42)	0.05	(1.12)
Education	8.57	(3.61)	8.50	(3.62)	8.61	(3.61)	-0.11	(-0.30)
Missing education	0.12	(0.33)	0.13	(0.33)	0.12	(0.32)	0.01	(0.23)
Children	0.77	(1.08)	0.71	(0.94)	0.80	(1.16)	-0.09	(-0.80)
Chg. married	0.03	(0.34)	0.05	(0.28)	0.02	(0.37)	0.04	(1.12)
Chg. children	0.83	(1.09)	0.99	(1.16)	0.73	(1.04)	0.25*	(2.19)
82-90 cohort	0.35	(0.48)	0.42	(0.50)	0.31	(0.46)	0.11*	(2.16)
91-00 cohort	0.44	(0.50)	0.39	(0.49)	0.47	(0.50)	-0.08	(-1.65)
Newspaper	2.26	(1.43)	1.96	(1.28)	2.47	(1.49)	-0.50*	(-2.14)
Saving share	0.06	(0.10)	0.07	(0.11)	0.05	(0.10)	0.02	(1.12)
Observations	394		150		244		394	

Notes: This table provides descriptive statistics for all migrants who (i) were at least 18 years old when they arrived in Germany, (ii) replied to SOEP within their first 2 years after arrival, (iii) replied for the first time at the latest in 2010 and (iv) replied to the return intention question. Compared to Table 2, this unconstrained sample also includes respondents who did not provide information required to construct our pessimism measures. [1] By design, the number of years of unexpected staying is only available for immigrants whose initial intention was to stay temporarily. Among them, the unexpected stayers have on average stayed 16 years longer than initially planned, while those who left Germany stayed 2.2 years longer than planned. . [2]This variable equals 1 for individuals for whom information on the number of education years is missing (15% of the sample) and 0 otherwise. For the latter, we assume their education years are equal to 0. Arrival cohorts 1982-90 and 91-2000 provide the share of respondents who arrived within a specific decade. The remaining 22% of respondents arrived between 2001 and 2010. ***,**,* denote significance at the 1, 5 and 10% level, respectively.

B.3 Ideal data

Based on our theoretical framework, we describe here the ideal data that would allow us to establish conclusively (i) that migrants underestimate their likelihood of staying and (ii) whether and how much pessimism over wage outcomes and over utility contribute to this underestimation. Since we focus on the ideal data we would like to have, we do not emphasize problems of implementation and measurement error here. Such issues are important, but as our data section illustrates, the primary problem with currently available data is that it requires strong identifying assumptions on preferences, homogeneity, and stability of migrant populations to offer an answer to these questions. We start by discussing the requirements for properly identifying mispredictions about migrants' long-term location, and then discuss the data needed to measure mispredictions about the distribution of future wages and about future utility.

B.3.1 Measuring unexpected staying

In order to properly assess whether migrants underestimate their likelihood of staying, the ideal data should provide probabilistic predictions to elicit the respondent's predictions about their probability of settling permanently in the host country. A growing literature shows that people are willing to provide probabilistic expectations in general contexts, and that these expectations contain useful information about future behavior - see Manski (2004) and Delavande (2014) for reviews in developed and developing countries, respectively. The questions, adapted to the context of migrants to Germany, would be of the form "What is the percent chance that you will settle in Germany permanently?".

A second requirement would be to follow migrants over time and to compare their subjective expectations ex ante to their actual location outcomes ex post. Obviously, the notion of ex post here suggests that a sufficiently long time has passed since migrants' arrival, so a long panel or survey data interconnected with administrative data is needed. Even so, in an ideal dataset, the migrant's final location would only be observed at death, though unexpected staying can take place at the intensive margin, i.e. when the migrant's actual duration of stay significantly exceeds their predicted duration of stay, a pattern that is observed in our data.

Another important aspect to assess imperfect predictions pertains to the inherent uncertainty faced by migrants. They may indeed experience events that they had not foreseen at arrival (e.g. falling in love with a native or another migrant, getting married or divorced, having children or not, losing their job, losing a parent in the origin country,...) that foster or deter a permanent stay. These shocks should ideally be observed and taken into account in the analysis, and one has to assume that the remaining unobserved shocks are balanced and independent across migrants. A correlated shock instead would be for instance if all migrants in the sample were from the same origin country and a civil war broke out in their home country, making all migrants more willing to

stay simultaneously.

If these conditions are met, and if migrants are observed in a sufficiently large number, one could conclude that if they have correctly calibrated (or 'rational') expectations, the predicted probability of staying permanently ex ante should on average be equal to the proportion of settled migrants ex post.

While the SOEP follows migrants since 1984, and many life events are captured, it does not capture migrants' predictions in the form of probabilities, the first requirement on ideal data that we mentioned. Instead, information on predictions relies on answers to the questions "Do you plan to return to your native country?"or "Do you want to stay in Germany forever?", which rounds probabilities of staying permanently in Germany to either 0 or 1. The most natural way of interpreting this question is to take the fraction of migrants who state that they intend to stay permanently in Germany as the percentage of migrants we expect to still be in Germany in the long term. The problem is that such a measure can be, in the most extreme case, off by up to 49% from the true subjective beliefs. To see this, consider homogenous migrants whose predicted likelihood of staying permanently in Germany is 49%. Since the survey only captures a binary prediction (permanent versus temporary stay), all these migrants would predict "yes, I plan to return to my native country at some point". Yet, 49% would stay permanently in Germany ex post. So the data would suggest unexpected staying whereas migrants' predictions are in fact correct, though imperfectly captured. Note that this does not mean that unexpected staying is systematically induced by this imperfect measurement. Consider the following counterexample: if all migrants had a correctly predicted probability of permanent stay of 51%, we would observe that they all predict a permanent stay, and would wrongly interpret that the 49% of migrants who left ex post are unexpected leavers instead.

The combination of these two extreme cases illustrates the main issue when collapsing probabilities into dummy predictions: it is difficult to precisely identify the nature of mispredictions. More specifically, if, as in our data, the proportion of migrants who predict a permanent stay is lower than the proportion of migrants who eventually stay, the most likely explanation is unexpected staying, but it need not always be the case. To illustrate this, suppose that the (imperfect) prediction we observe is that no migrant thinks their stay will be permanent, which means that their predicted probability of a permanent stay is comprised between 0% and 49%. Suppose now that, ex post, 40% are still in Germany. The most likely explanation in this case is unexpected staying: this is indeed the case if all migrants have a predicted probability of permanent stay comprised between 0% and 39%, which is lower than the actual proportion of stayers. However, this observation is also consistent with migrants whose predicted probability of permanent stay is comprised between 41% and 49%, in which case some migrants (between 1% and 9%) overestimated their probability of a permanent stay, making them unexpected leavers.

B.3.2 Measuring Misperceptions of Wages and Utilities

We now turn to the question of how to measure the role of mispredictions of wages and utilities in unexpected staying. Thus, the two broad misperceptions we consider pertain to the likelihood of future outcomes (wages in particular) and to their future utility given some future outcomes (wages, integration, savings,...).

As in the previous subsection, it is preferable to measure migrants' predictions by means of probabilities rather than binary predictions. In the case of wages, one would need to elicit migrants' beliefs over the likelihood that they will obtain various ranges of wages. Note that the way we elicit the expectations should be explicitly conditional on still being in Germany in the future, as we can never observe the counterfactual wage migrants would have earned in Germany when they have left the country. Again, using a long panel or administrative data, we would then compare these predictions to the migrants' actual wage a number of years later. If migrants have correct perceptions about their labor market prospects, the mean predicted wage should be close to the actual mean wage that they obtain.

Furthermore, while migrants' pessimism (or optimism) about their wage prospects in Germany is an important determinant of their long-term location, their beliefs about their alternative wage in the country of origin also matters. Clearly, if migrants' expectations about wages in the origin country are homogeneous (whether correct or not), our theory predicts that migrants who are the most relatively pessimistic about the host country are more likely to become unexpected stayers. Ideally though, one should also elicit the wage expectations at origin.

Since we are also interested in migrants' perception of their future utility, a natural way to do so is to ask questions about future life satisfaction: "On a scale from 1 to 10, what is the percent chance that your life satisfaction will be 0-2; 3-4; 5-6; 7-8; 9-10 in X years in Germany ". However, utility depends on several factors, including consumption and integration as formalized in our model. Because of this, we can elicit mispredictions about utility that are distinct from mispredictions about wage outcomes only if we ask migrants' predicted life satisfaction conditional on different wage scenarios. Ideally, one would thus condition questions about future life satisfaction on a number of different realizations of wage outcomes (assuming migrants properly predict how they will be integrated). The question should then be:

"If you were to earn [5 wage intervals] in X years in Germany, what is the percent chance that your life satisfaction will be [5 life satisfaction intervals] ? – While this approach is already demanding, it is not immune to other omitted factors influencing utility which could possibly bias our measure of pessimism.

B.4 Impact of pessimism on temporary intentions, remaining in Germany and unexpected staying

Table A5 provides results of linear regressions of the following equation:

$$Y_i = c + \beta \cdot Pessimism_{i,t_i^0} + \gamma \cdot X_{i,t_i^0} + \delta \cdot Cohort\left(t_i^0\right) + \eta \cdot Shocks_i + \epsilon_i,$$

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 2020 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 is still in Germany although they had expected to be a temporary migrant, and value 0 if they left Germany, or have correctly predicted to still be in Germany in 2020. 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 UnexpS dummy and it captures the idea that unexpected staying and unexpected leaving are opposite phenomena. In column (5), we use information, when available, to compare the number of intended years of stay to the realized spell for temporary migrants. More concretely, UnexpSY redefines immigrants who mentioned a temporary migration intention but are still in Germany as "other migrants" (i.e. not unexpected stayers) if their intended duration of stay has not been reached by the year 2020. These immigrants could still leave within their initially intended migration spell.

We did not opt for this measure as our benchmark definition for four reasons. First, the model does not explicit mispredictions about the duration of stay but rather mispredictions about the probability of permanent versus temporary migration. Second, the intended duration of stay (expressed in years in the data) is only available for a subset of migrants who state a temporary migration intention. Since migrants who stated a permanent intention were not asked about their intended duration of stay, an arbitrary imputation, for instance based on an average life expectancy, would be needed to create this variable for these respondents. Third, the intended duration of stay is missing among many migrants with temporary intentions, while our sample size is already limited. Finally, 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.

Table A5 shows that migrants who were pessimistic about their wage prospects did not form different intentions ex ante, but tend to stay more in Germany in the long term. As a result, the

Table A5: Intentions, final location, unexpected staying and Wage-Pessimism

	(1)		(2)		(3)		(4)		(5)	
	Temp. intentions		Still in Germany		UnexpS		UnexpSL		UnexpSY	
	b	se	b	se	b	se	b	se	b	se
W-pessimism	0.042	(0.08)	0.103	(0.07)	0.184**	(0.08)	0.158	(0.10)	0.138	(0.08)
Age	-0.008	(0.01)	-0.009	(0.01)	-0.011	(0.01)	-0.017	(0.01)	-0.008	(0.01)
Female	-0.052	(0.11)	0.075	(0.10)	-0.032	(0.11)	-0.004	(0.14)	-0.031	(0.11)
Married	0.023	(0.12)	0.165	(0.12)	0.182	(0.14)	0.365**	(0.17)	0.102	(0.14)
Education	0.029	(0.03)	0.024	(0.02)	0.012	(0.03)	0.046	(0.04)	0.005	(0.03)
Missing education	0.369	(0.32)	0.218	(0.28)	0.140	(0.31)	0.521	(0.39)	0.053	(0.32)
Children	0.148**	(0.07)	-0.080	(0.06)	-0.044	(0.07)	0.016	(0.09)	-0.033	(0.07)
82-90 cohort	-0.120	(0.16)	-0.245*	(0.15)	-0.212	(0.17)	-0.277	(0.21)	-0.188	(0.17)
91-00 cohort	-0.286*	(0.15)	-0.199	(0.14)	-0.305*	(0.15)	-0.385**	(0.19)	-0.261*	(0.16)
Chg. married			-0.020	(0.12)	0.161	(0.14)	0.278	(0.17)	0.148	(0.14)
Chg. children			0.102**	(0.05)	0.006	(0.06)	0.005	(0.07)	-0.006	(0.06)
Constant	0.641*	(0.35)	0.748**	(0.34)	0.736*	(0.39)	0.365	(0.48)	0.736*	(0.39)
Observations	105		105		105		105		105	

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 2020. "Unexpected stayers" (with=1) are migrants who report an initial intention to leave Germany but who stayed until 2020. ***,**,* denote significance at the 1, 5 and 10% level, respectively.

more migrants are pessimistic about wages, the more they tend to stay unexpectedly. Results are stable across alternative measures.

Table A6: Intentions, final location, unexpected staying and LS-Pessimism

	(1)		(2)		(3)		(4)		(5)	
	Temp. intentions		Still in Germany		UnexpS		UnexpSL		UnexpSY	
	b	se	b	se	b	se	b	se	b	se
LS-pessimism	0.045**	(0.02)	-0.006	(0.02)	0.040**	(0.02)	0.039*	(0.02)	0.025	(0.02)
Age	0.003	(0.00)	-0.000	(0.00)	0.004	(0.00)	0.002	(0.00)	0.004	(0.00)
Female	-0.130*	(0.08)	0.006	(0.06)	-0.105	(0.07)	-0.125	(0.09)	-0.074	(0.08)
Married	-0.093	(0.11)	0.192*	(0.10)	0.016	(0.13)	0.148	(0.15)	-0.005	(0.13)
Education	-0.018	(0.02)	0.053***	(0.02)	0.004	(0.02)	0.035	(0.03)	-0.001	(0.02)
Missing education	-0.066	(0.25)	0.467**	(0.19)	0.043	(0.24)	0.376	(0.30)	0.003	(0.24)
Children	0.003	(0.04)	-0.043	(0.03)	-0.036	(0.04)	-0.048	(0.05)	-0.032	(0.04)
82-90 cohort	0.164	(0.15)	0.088	(0.11)	0.218	(0.14)	0.257	(0.18)	0.259*	(0.15)
91-00 cohort	-0.028	(0.11)	-0.100	(0.08)	-0.031	(0.10)	-0.121	(0.12)	-0.003	(0.10)
Chg. married			0.002	(0.09)	0.091	(0.12)	0.075	(0.14)	0.091	(0.12)
Chg. children			0.002	(0.03)	0.022	(0.04)	-0.018	(0.05)	0.017	(0.04)
Constant	0.708**	(0.30)	0.215	(0.26)	0.268	(0.33)	-0.062	(0.40)	0.263	(0.33)
Observations	166		166		166		166		166	

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 2020. "Unexpected stayers" (with=1) are migrants who report an initial intention to leave Germany but who stayed until 2020. ***,**,* denote significance at the 1, 5 and 10% level, respectively.

Table A6 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 become unexpected stayers. Education and marriage are positively associated with staying ex post.[52] Tables A5 and A6 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 columns (4) and (5) has a slight impact on the significance of the pessimism measures, due to a minor decrease in the coefficient estimates combined with a slight increase in standard errors, but overall results remain stable.

Appendix B.5 shows how short-term decisions about integration and savings are linked to pessimism, temporary intentions and other covariates.

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

B.5 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 savings s are captured by the share of the migrant's monthly household income that is saved.

Table A7: Pessimism and short-term decisions

	(1)		(2)		(3)		(4)	
	Newspapers		Saving share		Newspapers		Saving share	
	b	se	b	se	b	se	b	se
W-pessimism	-0.310	(0.37)	-0.012	(0.02)				
LS-pessimism					-0.079	(0.10)	-0.002	(0.00)
Temp. intentions	-0.246	(0.39)	-0.030	(0.03)	0.075	(0.37)	0.015	(0.02)
Age	0.046*	(0.02)	-0.001	(0.00)	0.025	(0.02)	-0.000	(0.00)
Female	0.466	(0.48)	0.003	(0.03)	0.550	(0.38)	0.005	(0.02)
Married	-0.761	(0.51)	-0.057	(0.04)	-0.138	(0.68)	-0.018	(0.02)
Education	0.105	(0.11)	0.002	(0.01)	0.248**	(0.11)	0.005	(0.01)
Missing education^2.	0.658	(1.23)	0.039	(0.09)	1.616	(1.23)	0.048	(0.06)
Children	-0.130	(0.26)	0.013	(0.02)	0.003	(0.21)	0.001	(0.01)
82-90 cohort	1.195*	(0.68)	0.000	(0.09)	0.000	(.)	0.000	(.)
91-00 cohort	0.965*	(0.52)	-0.049	(0.04)	0.863*	(0.48)	-0.027	(0.02)
Constant	-0.187	(1.20)	0.170	(0.10)	-2.273	(1.51)	0.040	(0.07)
Observations	33		67		53		130	

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.

Table A7 shows that 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 significant link between pessimism and migrants' saving share either.