THE LONDON SCHOOL OF ECONOMICS AND
POLITICAL SCIENCE

Essays in Applied Microeconomics

Gianpaolo Caramellino

A thesis submitted to the Department of Economics of the London School of Economics for the degree of Doctor of Philosophy.

London, March 2018

Declaration

I certify that the thesis I have presented for examination for the PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).

I warrant that this authorization does not, to the best of my belief, infringe the rights of any third party.

I declare that my thesis consists of approximately 44,000 words.

Statement of conjoint work

I certify that chapters 2 and 3 of this thesis are co-authored. Chapter 2 is co-authored with Leonardo Felli, Carola Frege and Yona Rubinstein, while chapter 3 is co-authored with Francesco Sannino. I contributed 25% of the work in chapter 2, and 50% of the work in chapter 3.

Acknowledgements

I wish to thank several people that gave me fundamental inputs through my PhD.

First, I express my gratitude to my supervisors Francesco Nava, Leonardo Felli, and Yona Rubinstein. They constantly encouraged me during the hard times of the PhD by always being friendly and sympathetic. I truly believe that this thesis would not have been possible without their help. Furthermore, the exposure to their intellectual curiosity and rigor, often discussing topics that go well beyond these pages, will be beneficial also in my professional life.

I am grateful to other professors at the LSE, too. Carola Frege has been an extraordinary co-author, while Guy Michaels and Steve Pischke offered me many helpful comments at different stages of my work.

I acknowledge the financial support of the Economic and Social Research Council (ESRC) and the Fondazione Einaudi di Torino. At the Fondazione Einaudi, I thank Fabrizio Goria for his help and Giuseppe Bertola for his interest and comments on my research.

For their kindness and support during the transition between the undergraduate studies and the PhD, I am indebted to Dino Gerardi, Filippo Taddei and Pietro Garibaldi. I also thank Raffaello Bronzini and Silvia Magri, two great co-authors of mine, for being excellent guides during the initial phase of my research activity.

In London I met several new friends and I could rely on existing ones. I feel honored to have shared part of this journey with bright minds and nice people. I thank Alberto Pellicioli, Andreas Ek, Clare Balboni, Diego Battiston, Esteban Jaimovich, Fabio Pasquale, Giulia Bovini, Giulia Giupponi, Konstantinos Tokis, Marco Pelosi, Michel Dummar Azulai, Simone Fabiole-Nicoletto, Stephan Maurer, Sutanuka Roy, Thomas Drechsel, Tsogsag Nyamdavaa, and many others. I will miss the laps at Lincoln's Inn Fields and particularly those with Alexia Delfino, Davide Porcellacchia, Federico Rossi, and Nicola Limodio.

Special friends deserve special mentions. Alexandra Scudo has been almost like an additional sister in London. Sometimes I failed to be in touch, yet she has always offered me real breaks from the stress of my PhD, with inspiring chats and tasty dinners. Francesco Sannino and Matteo Benetton have been my greatest discoveries in London, and a great source of fun. I am indebted to Francesco, a sharp co-author, for always being thoughtful and supportive with me. I greatly benefited from his breathtaking ability to understand and address my needs, as well as from his intriguing viewpoints on several issues. Matteo, one of the kindest people I have ever met, is a true source of inspiration. In our long-lasting experience as officemates, he taught me how to practically pursue my dreams.

Finally, I am blessed to share the joy of my PhD with the people that are always on my side. Giulia was the main reason for doing my PhD in London and I will always be grateful to her for inducing me to take the right decision. I am thankful for her love and for her never ending support. Lalla inspires me every day with her examples and she has been truly important for me over these years. I feel extremely lucky to have her as my sister and I thank her for always being confident in my talent. My deepest gratitude goes to Mamma and Papà for their unconditional love and tireless support.

I wish I could share this achievement with Nonna Giovanna and Nonna del Maccagnotto and thank them for their love, cheerfulness, and examples of hard work.

My greatest academic achievement is for my most important teachers, my Mother and my Father.

Abstract

This thesis consists of three chapters that belong to the realm of Applied Microeconomics. The first two chapters are empirical projects that assess the role of time for human capital development of immigrants in the U.S.. The third one is a theory project that studies how managerial career concerns and experimentation influence risk-taking behaviours.

Chapter 1 studies how age at arrival in the U.S. affects the skill development of young immigrants in the U.S.. Using the within family variation across siblings entered in the U.S. at different ages, I document a cognitive / non-cognitive trade-off induced by age at arrival. As for cognitive skills, the effect of age at arrival is negative, in particular for the ability to learn English. The effect on cognitive skills is reflected in immigrants' educational achievements. However, age at arrival plays a positive role for illicit behaviours. Children of immigrants arrived later tend to show less problematic behaviours than their siblings arrived earlier, also controlling for their English ability. Through an indirect accounting exercise, I estimate the negative effect of age at arrival on the labor market performance of immigrant adults. I conclude the paper showing that more educated parents anticipate the arrival of their children in the U.S..

Chapter 2, co-authored with Leonardo Felli, Carola Frege and Yona Rubinstein, studies the intergenerational assimilation of immigrants in the U.S.. In our study, we observe the outcomes of several immigrant generations. Moreover, we link immigrant mothers and their children, thus observing the outcomes of two immigrant generations belonging to the same cohort. Controlling for the selection into migration and return migration, we document that it takes two immigrant generations to exhaust the full potential of cognitive and educational assimilation, while it might take longer for other social outcomes, such as the attitude towards problematic behaviour and the likelihood of having children.

Chapter 3, co-authored with Francesco Sannino, studies the effect of managerial career concerns and experimentation on risk-taking. We model an economy where managers create value through their ability to learn at an intermediate stage about the intrinsic profitability of a risky investment. Managers are heterogeneous in their ability to extract information from experiments, and care about their reputation. Their incentive to take on risk is distorted by career concerns, and can result in under or over risk-taking. When, following the experiment, better managers discard risky projects more often than bad ones we observe over risk-taking. Our result is in contrast with Holmström (1999) where managers' ability affects the project's success rate, and career concerns can only produce inefficiently low risk-taking. We show that the inefficiency is reduced in one extension of the model, where the market can also observe the outcome of similar projects. The novel implication is that markets more plagued by career concerns distortions are those where managers engage in more idiosyncratic activities.

Contents

List of Figures

List of Tables

Chapter 1

Do not Put Off Until Tomorrow What You Can Do Today: Age at Arrival and Immigrants' Human Capital

1.1 Introduction

The number of immigrants and children of immigrants in most of the richest countries is rising. With the increase in international mobility, individuals that face different incentives and additional barriers in the accumulation of skills are more and more important in determining these nations' stock of human capital and their relative advantages compared with each other. Children of immigrants, relocating around the critical periods for the full development of their personalities, play a particular role in this process. Although they do not choose to become immigrants, they have to adapt their human capital investment process to the new educational environment and, on average, they bear the cost of migration for longer.

Is age at arrival an important factor in explaining the pre-labour market outcomes of children of immigrants in the U.S.? Is a late arrival unambiguously bad? These are the issues that I take up in this paper. In trying to answer these questions, I make use of the NLSY79 dasaset, that allows me to shed light on the differential role of age at arrival in shaping several dimensions of individuals' personality, often considered unobservable. In the pages that follow I show that age at arrival has, in general, a negative effect on the educational achievement of children of immigrants, that is also reflected in the performance in standardized cognitive tests. It has, however, a positive effect on the likelihood of not being involved in anti-social and illicit activities. I also provide indirect evidence that the net effect of this trade-off is tilted toward a preference for arriving earlier.

From an econometric perspective, assessing the impact of age at arrival on human capital accumulation poses the challenges associated with the endogeneity of the former. There are several reasons for why age at arrival should not be considered as good as randomly assigned. First, certain households' characteristics can influence the timing of the migration. For example, some parents might understand better than others the problems induced by switching from the schooling system or the institutional environment of the home country to the American ones, and thus they might be tempted to move to the U.S. as soon as possible. Similarly, families that have the capabilities to mitigate the cost of a relocation or that have less children are likely those with the material resources to pick the timing of migration from a larger opportunity set. Moreover, while immigrant parents relocating to a new country for labour related reasons are often able to choose the timing of the migration, other people leaving their home countries for other reasons, such as refugees, are not. Because some or all these factors are also relevant in the individuals' human capital production function, ordinary least squares capture, in part, spurious correlations between the outcomes and age at arrival. One of the advantages of the

NLSY79 is that it offers the possibility of observing immigrants whose age at arrival is different, but that are instead similar along the other dimensions that affect their human capital. Specifically, to assess the role of age at arrival on the human capital development of immigrants, I use family fixed effects and I exploit the variation in age at arrival across different siblings within the same household.

Another advantage of the data I use is the richness of individuals' information I can observe. A part from the numbers of years of education, I am able to study immigrants' performance in standardized cognitive tests across several subject areas[1], as well as to understand their social behaviour. I find that an extra year in the home country reduces the number of years of education by 0.14 to 0.2, with the critical period around the age of 10. As for cognitive skills, my results indicate that the degree to which a late arrival is detrimental is subject-specific. In particular, the effect is stronger for the knowledge of English language, and milder in case of mathematics. Furthermore, my results indicate that reducing the age at arrival by one year increases the amount of illicit, risk-taking, and problematic behaviours by about 0.05 to 0.15 of a standard deviation. Despite this apparent trade-off in the development of cognitive and non-cognitive skills, I then provide suggestive evidence that an earlier arrival is preferable compared to a later one, as it leads to better labour-market outcomes. I also show that parents with higher education bring their children earlier.

Researchers have hinted at several possibilities for explaining the role of age at arrival in the assimilation process of children of immigrants. The inability in acquiring language proficiency after a certain threshold age[2], tied to the importance of verbal ability in acquiring non-linguistic skills or in attaining success in school, is

[1] Other papers proposed the performance in these tests as important predictors of labour market outcomes. See, for example Heckman et al. (2006).

[2] This hypothesis builds on the psychological literature on the *critical period hypothesis*. The seminal paper in this literature is Lenneberg (1967), whereas, more specifically, Johnson and Newport (1989) studies the critical period for learning a second language.

conceivably an important one. Although I lack the exogenous variation in English knowledge to run the ideal experiment and to assess its causal effect as a channel through which age at arrival has an impact on non-verbal outcomes, when I use verbal ability as an additional controller in my regressions of schooling and cognitive skills on age at arrival, the estimates reduce substantially. I rule out, instead, any role of linguistic knowledge in determining the relation between age at arrival and the likelihood of being involved in illicit behaviours.

Relation to the literature. This paper contributes to the debate on the role of age at arrival on the immigrants' human capital formation and their socio-economic performance. Several papers study the effect of age at migration on academic achievements; examples are Gonzalez (2003), Bleakley and Chin (2004), Van Ours and Veenman (2006), Ohinata and van Ours (2012), Böhlmark (2008), Basu (2016) and Lemmermann and Riphahn (2017). While the studies based on U.S. data (Gonzalez (2003) and Bleakley and Chin (2004)) do not account for the endogeneity of age at arrival (with the exception of Basu (2016)) and use the number of years of education as the main outcome of interest, I use family fixed effects and a broader set of measures of cognitive skills. From a methodological perspective, my paper is similar to those based on Swedish and German data (Böhlmark (2008) for Sweden and Lemmermann and Riphahn (2017) for Germany). In this respect, my paper adds to these studies because the characteristics of the host country and its selection of immigrants are different[3] and because I also consider some non-cognitive skills.

My paper also speaks to the literature on the importance of verbal knowledge for educational and economic achievements. Bleakley and Chin (2004) show that English ability is an important predictor in wage regressions of immigrants and the

[3]For example, the estimates for the effect of age at arrival on years of education in Lemmermann and Riphahn (2017) are about one third of those in Gonzalez (2003) and Bleakley and Chin (2004), which are closer to my estimates. This may suggest that the age of arrival effect on educational achievements is heterogeneous across countries.

effect is mainly due to its role in affecting immigrants' education. Isphording et al. (2016) finds that increasing linguistic ability by one standard deviation increases the math performance in standardized tests by 0.6 of a standard deviation. Aucejo et al. (2016) conclude that linguistic ability is more important than math skills in determining university enrolment and that teenage years are particularly important for developing skills. In my paper I find a strong negative effect of age at arrival on the ability to learn English. In the context of this literature, my result implies that reducing children of immigrants' age at arrival, by increasing their English ability, should be beneficial for the development of their non-linguistic skills and for their labor market outcomes. This result is in contrast to Lemmermann and Riphahn (2017), that using German data find that linguistic knowledge is not a relevant channel through which age at arrival affects educational achievements.

More broadly, my paper is related to the studies documenting gaps between first generation immigrants and natives (see for example Chiswick and DebBurman (2004) and Algan et al. (2010)), as well as to the literature on the cognitive and non-cognitive skill formation of young individuals and on early childhood intervention (see Cunha and Heckman (2007) and Cunha et al. (2010)). As such, my study indicates that providing incentives to immigrant parents to anticipate their arrival in the host country would reduce the cognitive barriers their children face in investing in human capital, thus reducing the gap between them and the children of native parents.

The rest of the paper is organized as follows. Section 1.2 describes the econometric issue and the solution used in this study. Section 1.3 describes the data. Section 1.4 presents the main results. Section 1.5 studies the possible mechanism of the main effects, addresses the policy trade-off induced by age at arrival and studies the parental selection. Section 1.6 reports some robustness checks. Section 1.7 concludes.

1.2 Empirical strategy

In order to guide the empirical analysis, I start by developing a simple conceptual framework that describes different determinants of immigrant children's age at arrival. A toy model formalizing the discussion can be found in the Appendix. I then show why in my empirical model the endogeneity of age at arrival makes the estimation problematic and how I can address this problem using family fixed effects and the NLSY79 data.

Conceptual framework

The sample in my study consists of immigrant children arrived in the U.S. when they were teenagers or younger. Presumably, for most of them, the decision on when to arrive was their parents' choice. For simplicitly, let us think of the economy as a two-period world. In the first period, immigrant parents choose when to move to the U.S.. In the second one, immigrant children observe their age at arrival and choose their investment in human capital, accordingly. The cross-sectional empirical relation between the investment in human capital and age at arrival is the combination of an age at arrival effect and a selection one, that arises because of common factors affecting both parents' and children's optimal decisions.

To understand where the age at arrival effect comes from, suppose, for the moment, that the selection effect is null. Immigrant children set the level of investment in human capital so that the marginal benefit they get from it equals its marginal cost, conditionally on age at arrival in the U.S.. Assuming that children's problem is concave and, hence, the solution is interior, the sign of the age at arrival effect is determined by how the marginal cost of human capital, relative to its marginal benefit, changes with age at arrival. For example, the literature suggests that age at arrival might increase the cost of investing in U.S.-specific skills, through its effect on linguistic proficiency (see the seminal paper Lenneberg (1967) on the critical period hypothesis to acquire linguistic proficiency). On the other hand, certain aspects of the new local environment might as well be detrimental for the future individual's success. In this case, a higher age at arrival could be beneficial for children of immigrants, by increasing the likelihood of retaining certain positive cultural traits typical of their home countries.

The selection effect, instead, suggests that age at arrival in the U.S. is likely not as good as randomly assigned. Immigrant children that arrived earlier may be systematically different from those arrived when they were older, in terms of attitudes towards skills accumulation or costs in acquiring them. There are several reasons that support this possibility. Parents select the timing of migration, and hence the age at arrival of their children, depending on their own preferences, abilities or material resources. Furthermore, some families might be forced to leave their home countries without the possibility to select the best timing. Those family characteristics might, in turn, be correlated with some determinants of children's investment in skills.

Consider parental preferences and abilities, first. Some parents could value their children's U.S.-specific skills more or anticipate the age at arrival effect better than others. They could then choose the optimal timing of arrival in the U.S. to help their children get those skills. As preferences, beliefs and norms determining both parents and children's behaviours are, at least partly, transmitted by genes or through personal interactions (see Bisin and Verdier (2011)), these characteristics would likely be correlated with some determinants of their children's investment in human capital. Material resources could play an important role, too. Certain families may be able to reduce the impact of age at migration by selecting wealthier neighbourhoods or exclusive schools for their children. These families are also likely those that can pick the timing of migration from a larger opportunity set. A similar possibility is that larger family size, whose effect on achievements is negative (see, for example, Black et al. (2005)), increases the cost of arriving earlier[4]. In general, individuals arrived earlier are not necessarily those more willing to invest in U.S.-specific human capital. Immigrants enter in the U.S. for different purposes and some immigrant categories, particularly refugees or asylees, are less likely to freely choose the timing of migration. Simultaneously, some of them may be more easily eligible for permanent residence or naturalization (see, for example, Woodrow-Lafield et al. (2004)) or less likely than others to migrate back to their origin countries, and thus possibly more willing to invest in U.S-specific human capital.

Econometric model

The objective of my study is to estimate the coefficient β_a in the regression

$$y_{it} = cons + \beta_a a_{it} + \beta_x X_{it} + \epsilon_{it}$$

where y_{it} is a measure of skills of individual t living in household i, while a_{it} represents his age at arrival in the U.S.. X_{it} is a set of characteristics that may vary at the individual level and affect the skill accumulation, and ϵ_{it} is an idiosyncratic error term.

In the previous subsection I explained some reasons for why a_{it} and the shock ϵ_{it} might be correlated. If these concerns are valid, ordinary least squares estimates that make use of the whole cross-sectional variation in age at arrival would provide biased estimates of the causal effect of age at arrival on skill accumulation, β_a. In this subsection I explore what I can do, as an empirical researcher, using fixed effect regressions and the NLSY79 dataset.

[4]For example, grandparents' help in the home country may be an increasing function of the number of grandchildren. However, as they do not relocate to the U.S., it would drop to zero after migration.

In an ideal experiment I would randomly select immigrant children in the sample and force them to enter the U.S. at different ages. In this way, there would be no systematic link between the age at arrival in the U.S., a_{it}, and the realization of the shock ϵ_{it}. As such an experiment is impossible, I use a model with family fixed effects,

$$y_{it} = cons + \beta_a a_{it} + \beta_x X_{it} + \overbrace{\gamma_i + \gamma_{it}}^{\epsilon_{it}}$$

where I assume that the shock ϵ_{it} has two components: γ_i is a fixed family component, whereas γ_{it} is the idiosyncratic individual component.

The critical assumption to estimate consistently the causal effect of age at arrival, β_a, using this fixed effect specification is that γ_i is the only reason for the correlation between ϵ_{it} and a_{it}. That is, conditional on having the same parents, age at arrival in the U.S. is as good as randomly assigned for pairs of siblings. Two additional pitfalls of this strategy might nonetheless affect the consistency of the family fixed effects estimator. First, if immigrant children, rather than their parents, can choose or, at least, can influence their age at migration, their decision may be correlated with some determinants of their human capital production function. In order to address this concern, I only focus on individuals that arrived at the age of 17 or younger. Second, immigrant parents may treat differentially different children. For example, they might prefer one child over the others. In some cases, they may allow the year of migration to vary across children within the same household, depending on some characteristic that are child-specific, but unobservable to the econometrician. For about 1/6 of the immigrant children in my sample, the year of arrival in the U.S. does not match the one of their siblings. I therefore run the regressions for both the whole sample and the sample restricted to siblings arrived in the same calendar year. Furthermore, even if the year of arrival is the same, parents can still choose it to maximize the achievements of only a specific subset of their children. To partially address this concern, I always control for sex and, when specified, for a birth order indicator.

When I use the outcomes in standardized tests to measure the cognitive skill accumulation of children of immigrants, as well as when I consider attitudes toward illicit behaviours, I also face an additional problem. The concern, in this case, is the perfect collinearity between age at arrival and age at the time of the interview. In the NLSY79, indeed, individuals take the cognitive tests or answer to behavioural questions in the same year, but at different ages. Potentially, both these two components have an effect in determining the outcomes. To solve this problem, when the outcome variable of the regression is the performance in cognitive tests, or the attitude towards illicit behaviours, I add natives to my sample. By assuming identical cohort effects for natives and immigrants, the variation across native siblings identifies the age effect, while, once the cohort effect is partialled out, the residual variation across immigrant siblings identifies the age at arrival effect on the outcome of interest.

To conclude this section, I also want to stress that measurement error in age at arrival is likely to occur in this study. As some individuals are asked several years later about the year in which they entered in the U.S., some of them could report it wrongly. The fixed effect estimates exacerbate this problem. If the conditions for the classical measurement error apply to my analysis, my estimates would be biased toward 0[5]. It is important to notice, however, that restricting the sample to siblings arrived in the same calendar year should make the measurement error less severe.

[5] See Angrist and Pischke (2008).

1.3 Data

Data sources and variables of interest

In my study I use the National Longitudinal Survey of Youth 1979 (NLSY79) and the U.S. Census.

The NLSY79, that is the basis for the main part of this study, is a survey administered by the U.S. Bureau of Labor Statistics to 12,686 individuals living in the U.S.. The survey participants were born during the period 1957 through 1964 and have been interviewed annually since 1979 through 1994 and, since then, on a biennial basis. The NLSY79 has been intensively used in the labor literature, because it offers detailed information on several socioeconomic outcomes of a representative sample of the youth population in the U.S.. From the NLSY79, I obtain the individuals' number of years of education, the outcomes in standardized cognitive tests and an index of illicit behaviours, as well as information on age at arrival in the U.S. and family structure.

To measure the development of cognitive skills, in addition to educational achievements, I use the results in several sections of the Armed Services Vocational Aptitude Battery (ASVAB) test. The ASVAB is the test administered by the U.S. Military Entrance Processing Command to assess the cognitive skills of people that want to join the army. NLSY79 respondents took the ASVAB test because the U.S. Departments of Defense and Military Services wanted to update the norms of the ASVAB, using nationally representative samples of young people. Each participant in the NLSY79 took the test in 1980, although results are missing for the 6% of the original 1979 sample. In this study, I focus on the performance in the three mathematics and in the two English sections of the ASVAB. The sum of the scores in two mathematics sections, the arithmetic reasoning and the mathematics knowledge sections and in two English sections, the word knowledge and the paragraph comprehension ones, determines the Armed Forces Qualification Test (AFQT). The AFQT has been often used in the literature as a proxy of IQ (see, for example, Cameron and Heckman (1998), Cameron and Heckman (2001), and Heckman et al. (2006)). In the arithmetic reasoning part, candidates face math word problems and the goal of these questions is to assess their ability to apply mathematics in solving real world problems[6]. The mathematics knowledge section covers basic high school mathematics, including questions on algebra and geometry. The numerical operation part of the test is a speed test in performing simple mathematical computations. In the word knowledge section candidates must pick the best synonym of a given word in a sentence[7]. In the paragraph comprehension one, candidates are evaluated on their ability to understand brief reading passages[8]. My measures of linguistic knowledge differs substantially from other measures used in the immigration literature, that are often self-reported and expressed as categorical variables with few potential values. In my analysis, I standardize the cognitive skill measures to have a mean of zero and a standard deviation of one in the NLSY79 population.

The index of illicit behaviours is useful to obtain a glance at the development of a particular sociability trait of immigrants in the U.S., that is often a motive of debate, although difficult to measure. I construct an index which is computed using questions from the 1980 survey, and measuring the degree to which an individual engages in aggressive, risk-taking, and illicit behaviors. Each survey participant was asked to answer to twenty questions: seventeen are about delinquency and

[6]For example, candidates must solve the problem: "One in every 9 people in a town vote for party A. All others vote for party B. How many people vote for party B in a town of 810?"

[7]As an example, candidates must choose the most appropriate synonym of *unblemished* in the sentence "The employee was proud of her *unblemished* reputation," out of *impaired*, *flawless*, *sporadic*, *unorthodox*.

[8]For example, respondents must read the passage "The Mississippi River is key to New Orleans' flavor and pizzazz. The seafood, the steamboat cruise, the swamp tours, and the history-it's all there. And the jazz? There are those who would swear that the uncanny beat of the music comes from the intrepid rhythm of the Mississippi's waters." Then they have to pick the element that mainly affects the atmosphere and reputation of New Orleans between *jazz music*, *its unique food*, *the Mississippi River*, *the history of its swamp*.

three relates to problems with the police. To be more specific, the topics of these questions include property damaging, fighting with classmates, shoplifting, robbery, drug use, drugs dealing, charges or conviction by the police with an illegal activity. Following Heckman and Rubinstein (2001), I assign for each question a value of one if the person engaged in that activity and zero otherwise, I add these values and divide by twenty. I then standardize the index, which is available for about 90% of the sample, to have a mean of zero and a standard deviation of one in the NLSY79 population.

Importantly, immigrants in the NLSY79 were asked about their year of arrival in the U.S. in two survey years, in 1983 and in 1990. In the main part of the analysis, if both answers are non missing but are different, I use the average of the two as year of arrival in the U.S.. If one of the two is missing, I only use the remaining one.

To study the relation between age at arrival of immigrants in the U.S. and their parents' characteristics, I also use the U.S. Census data 1970-2000. These represent random samples of the American population and the sample size is consistently larger than in the NLSY79. It is important to notice that for Census years 1970 to 1990, the year of arrival (and, therefore, also the constructed age at arrival) is reported in intervals. In Census 2000, instead, people were asked about the exact year of entry in the U.S..

NLSY79 sample selection and summary statistics

In the NLSY79, I restrict the sample to individuals born abroad with both parents and the paternal grandfather[9] born abroad. I further limit the analysis to those arrived in the host country when they were seventeen years old or younger, the sometimes called 1.5 immigrant generation. First, these individuals likely did not

[9]The only grandparent with available information on country of birth in the NLSY79 is the paternal grandfather.

choose to migrate to the U.S.. For them, it is less likely that personality traits are part of the decision to become an immigrant, and, in particular, to arrive at a specific age in the U.S., rather than a consequence of that. Importantly, this choice also allows me to focus on people for which human capital investments in the host country are in part a matter of choice. Most U.S. states, in fact, have compulsory schooling laws requiring to attend school until at least the age of sixteen.

There are 537 individuals in the NLSY79 that satisfy these sampling requirements. For 256 of them, that live in 113 distinct households, I can identify at least one sibling in the survey. However, only 213 of them, corresponding to 94 families, declare a year of arrival that matches the one indicated by their sibling(s).

Table 1.1 reports the summary statistics of the main variables of interest for natives and immigrants in the NLSY79. As for immigrants, I report the summary statistics for the whole sample and for the sample of siblings. As expected, immigrants have worse educational achievements and get lower scores in cognitive tests. The gap in schooling between the average American and the average first generation immigrant is about 0.5 years. The gap in cognitive skills ranges between 0.3 and 0.8 of a standard deviation, depending on the subject. Perhaps not surprisingly, it is smaller for mathematics and larger for English ability. Interestingly, first generation immigrants behave better than natives, by about 0.35 of a standard deviation. The average age at arrival in the U.S. for first generation immigrants in the NLSY79 is just below the age of 10.

1.4 Results

Graphical analysis

Figure 1-1, which uses the whole immigrant sample in the NLSY79, offers the possibility of a first glance at much of the subsequent results.

There is a strong negative correlation between age at arrival and educational outcomes. An extra unit in age at arrival corresponds to a decrease in completed education by 1/4 of a year. The mode in education for most ages of arrival is 12 years. However, for people that arrived at the age of 4 or later it is not uncommon to attend school for less than 10 years, and for individuals arrived after the age of 9 it is quite frequent to complete no more than 6 years of education[10]. Furthermore, although there are individuals achieving the highest grades for almost any age at arrival, the relative frequency decreases with age at migration.

The results in educational achievements also reflect those in performance in cognitive tests. One extra unit in age at arrival decreases the performance in mathematics tests by 0.077 of a standard deviation, while it reduces the performance in English tests by 0.103 of a standard deviation. For both measures, the lowest scores are obtained by individuals arrived at the age of 10 or later. However, although the relative frequency appears decreasing with age at arrival, there are individuals able to perform better than the average individual in the NLSY79 even among immigrant children arrived after the age of 10.

In terms of probability of being involved in illicit activities, instead, age at arrival plays a positive role. The relation between the two is negative, meaning that the later an immigrant enters in the host country the less he acts illicitly. In particular, the correlation is driven by the almost complete absence of individuals that, arriving after the age of 11, behave more than a standard deviation worse than the mean individual in the NLSY79. It is also evident from the picture that the share of individuals that behave perfectly is an increasing function of age at arrival.

Similar patterns emerge when I use the sample consisting of immigrant siblings, reported in Figure 1-2. Apart from illicit behaviours, however, the correlation be-

[10]Some of these people were, potentially, drop out in their home countries since some years before their migration to the U.S.. This correlation, therefore, mixes two effects: the probability of becoming a drop out after the migration to the U.S., and the probability of not re-enrolling in school in the U.S. after dropping out in the home country.

tween the various outcomes and age at arrival are, in this case, reduced by 20% to 40%. It is also worth mentioning that the correlation between years of education and age at arrival is similar in the NLSY79 and in the Census dataset, as shown in Table 1.2. In the Census one extra unit in age at arrival is associated to a reduction in educational achievement by between 0.19 and 0.23 of a year, compared to a reduction between 0.14 and 0.26 in the NLSY79.

Main regression results

In Tables 1.3 to 1.9 I show the results of age at arrival on the accumulation of skills using the econometric model described before.

First, in Table 1.3 the results on educational achievements are reported. The OLS estimate in column 1 suggests that an extra year in age at arrival lowers the educational achievement by 0.14 years of schooling. In column 3, the result of the model including family fixed effects is identical. It is possible that my estimates are affected by the systematic parental preference for children born first. The estimated parameter, however, remains quite similar with the inclusion of controls for the birth order of individuals, as shown in column 5. In the latter case, however, the precision of the estimate reduces and the significance disappears. When I estimate a more flexible model with age at arrival defined through five categorical variables, whose results are reported in even columns, the fixed effects model suggests a critical age around 10. That is, individuals arrived around the age of 10 suffer a gap of about 0.9 years in completed education in comparison to their siblings arrived when they were 0 to 5, while for those arrived between the age 15 and 17 the gap reaches 1.6 years. In this case, the result is robust to the inclusion of birth order controls. Restricting the sample to siblings arrived in the U.S. in the same calendar year does not alter much the qualitative insights. The results reported in columns 7 to 12 indicate that the point estimates increase slightly.

As for cognitive skills, the results are in Tables 1.4 to 1.8[11]. Similarly to educational achievement, age at arrival has a negative effect on cognitive skills. The results of the linear models estimated with OLS regressions or with family fixed effects without controls for birth order, reported in columns 1 and 3, show that an extra year in age at entry in the U.S. lowers the measured outcomes by 0.02 to 0.07 of a standard deviation. Importantly, the effect is lower and non significantly different from zero for cognitive mathematical achievements, whereas it is stronger and significant for English knowledge. As for English knowledge, the effect on paragraph comprehensions is more important than the one on word knowledge. These conclusions are unaffected by the inclusion of birth order controls, as shown in columns 5. When I use a more flexible model with categorical ages at arrival and individuals arrived between the age of 0 and 5 as benchmark group, the gap, increasing in age at arrival, is significant for immigrants arrived after the age of 11. Also in this case the effect is stronger for paragraph comprehension, reaching, in the specification with all the controls, a difference of more than 1 standard deviation between individuals arrived around the age of 16 and those arrived before the age of 6. As for mathematics, it is instead lower and often very noisy, although significant for individuals arrived around the age of 13 in different specifications for each outcome. Once again, the results on both English knowledge and mathematics proficiency do not change much when I restrict the sample to siblings arrived in the same calendar year.

The results on illicit behaviour are shown in Table 1.9. The linear model estimated with OLS in column 1 shows that anticipating the arrival in the U.S. by one year reduces the probability of being involved in illicit activities by a statistically significant 0.05 of a standard deviation. Although the estimates of the fixed

[11]Notice that in these tables, as well as in the table on illicit behaviours, the number of observations is much higher, because I use also native siblings in the sample to identify the effect of age on the different outcomes.

effects models in columns 3 and 5 are not significantly different from zero, the point estimates are larger, reaching the level of 0.09 with the full set of controls. The more flexible model using family fixed effects, whose results are shown in columns 4 and 6, suggests that individuals arrived around the age of 7 show substantially better behaviours than those arrived before the age of 6, by about 0.5 of a standard deviation less problematic. Without controls for birth order, however, the results are similar to the ones obtained with OLS, but not significant. With the inclusion of birth order controls, instead, the difference in attitude toward behavioural problems becomes significant for individuals arrived around the age of 8 and around the age of 16. Restricting the sample to siblings arrived in the same year makes the estimates bigger in absolute terms and more often significant.

1.5 Further considerations

In this section I add further considerations to the the main results. I first check if age at arrival has an effect on education, skills accumulation and illicit behaviours through its effect on English proficiency. I then use an accounting exercise to estimate the effect of the trade-off between cognitive skills and illicit behaviours on labour market outcomes. Third, I study parental selection in age at arrival.

Possible mechanism

One possibility for my results is that with its negative effect on linguistic ability, age at arrival increases the costs or limits the potential to invest in the development of human capital. Although English ability is an endogenous variable in this context, I re-estimate the main regressions controlling for word knowledge and paragraph comprehension. The results are shown in Tables 1.10 and 1.11. First, in general, word knowledge and paragraph comprehension are strongly correlated with the educational and cognitive outcomes, while only paragraph comprehension is marginally correlated with attitudes toward illicit behaviours. Second, the point estimates of the age-at-arrival effect on education and accumulation of cognitive skills reduce substantially, whereas the estimates on illicit behaviours are almost identical to those in the main specification reported in Table 1.9. This means that, differently from Lemmermann and Riphahn (2017), I cannot reject the possibility that verbal ability is one of the channels through which age at arrival affects the development of cognitive skills. Instead, it does not seem to play any role in explaining the effect on illicit behaviours.

An accounting exercise

Overall, my estimates in Section 1.4 suggest that age at arrival is detrimental for educational achievements and for the accumulation of cognitive skills. It reduces, however, the attitude toward illicit behaviours. This implies a potential trade-off in designing migration policies aimed at attracting young children of immigrants.

To address this potential trade-off, I quantify the effect of an extra-year in the home country on labor market outcomes of immigrants in the U.S.. The small sample size, together with the attrition in the participation to the NLSY79 survey, prevents me to directly estimate the effect of age at arrival on labour market outcomes. I then proceed in two steps. In the first one, I estimate the following Mincerian equation

$$y_t = cons + \beta_e e_t + \beta_c c_t + \beta_i i_t + \gamma_1 exp_t + \gamma_2 exp_t^2 + \gamma_X X_t + \epsilon_t$$

where y_t represents the labour market outcome of interest of individual t, such as the likelihood of being employed and the wage. e_t stays for educational achievements, c_t indicates a full set of linguistic and mathematical cognitive skills, i_t stays for a measure of illicit behaviours, while exp_t is the potential working experience of the individual. In the set of controls X_t, I include gender and ethnicity. In the second step, I use the estimates of β_e, β_c and β_i, together with the estimates of the previous section, to approximate the labour market impact of an extra unit in age at arrival, as follows,

$$\widehat{\frac{dy}{da}} = \widehat{\beta}_e \widehat{\frac{de}{da}} + \widehat{\beta}_c \widehat{\frac{dc}{da}} + \widehat{\beta}_i \widehat{\frac{di}{da}}$$

The various $\widehat{\beta}$ are the estimates from the first step of my accounting exercise, reported in Table A.22, while $\widehat{\frac{de}{da}}$, $\widehat{\frac{dc}{da}}$ and $\widehat{\frac{di}{da}}$ are the partial effects of age at arrival on education achievements, cognitive skills and illicit behaviour, respectively, estimated in Section 1.4.

The results of this exercise indicate that an extra unit in age at arrival reduces the probability of being employed by less than 0.003, while, conditional on being employed, the likelihood of being employed full time by 0.0025. Both the hourly wage and the yearly earnings decrease by about 0.02 as age at arrival increases by one unit. The results are very similar when I use the much larger sample of American in estimating the Mincerian equation. The only difference is the estimated likelihood of being employed full time, close to zero in this case.

In any case, it is important to notice that if policymakers interpret immigrants' assimilation as looking like an American the trade-off is non-existent. As Table 1.1 shows, indeed, first generation immigrants tend to behave better than natives and my results suggest that by arriving earlier immigrant children would be more similar to the average American individual along all the human capital dimensions.

Parental selection

I now look at the parental selection in age at arrival. If parents care similarly to the human capital development of their children and the net age at arrival effect is negative, those that possess the material or intellectual resources to move earlier to the U.S. should do so.

The results are shown in Table 1.12. Using the whole immigrant sample in the NLSY79, an extra year in mother's education is associated to a reduction in age at arrival by more than 0.1 units, while the effect of father's education is a reduction by 0.03. I obtain similar results focusing on families with two or more siblings arrived in the U.S. as children of immigrants. In this case, the effect of father's education increases in absolute value, being associated to a reduction of 0.17 units in age at arrival. The coefficients are, however, not statistically different from zero. Interestingly, when I also control for family income, I find that the correlation between parental education and age at arrival is basically unchanged, while family income has no role in explaining the age at arrival of children in the U.S.. Since in the NLSY79 I observe family income only in the 1979, I re-estimate in column 4 the same model focusing on immigrants not yet in working age in 1979, and my conclusion does not change. In this case, however, the effect of maternal education on age at arrival of children is smaller in absolute terms.

The findings in this section suggest that parents move to the U.S. taking into account the age at arrival of their children. The selection effect in theory might be driven by material resources allowing to choose the timing of migration from a larger opportunity set, or by parental intellectual ability to anticipate the age at arrival effect on children' cognitive and non-cognitive development. Although my exercise is limited, my findings suggest that the second element is likely more important than material resources.

To validate the results on parental selection, I also use the Census data, that provide a much larger sample of immigrant children. When I use the Census data, the estimates are qualitatively similar, but the larger sample size reduces the standard errors and the results becomes significant. In this case I find that mother's education decreases the age at arrival by around 0.12 units, while father's education by 0.07. These estimates are robust to the inclusion of cohort of arrival fixed effects and to the exclusion of Census years 1970 to 1990, when age at arrival was defined by intervals.

Heterogenity

In the following paragraphs I explore the extent of heterogeneity in the age at arrival effect, depending on the country of origin of immigrant children. Due to the small sample size most of the heterogeneity effects that I describe below are not statistically different from 0. Nonetheless, I believe it is useful to report these results.

In Tables A.14 to A.17, I interact age at arrival with an indicator for immigrants of Mexican origin. In the immigrant siblings sample of the NLSY79 there are 108 individuals arriving from Mexico and 148 arriving from other countries. In Tables A.18 to A.21, instead, I interact age at arrival with an indicator for immigrants arriving from countries where English is an official language. Unfortunately, however, there are only 28 immigrant individuals in the immigrant siblings sample that arrive from countries where English is an official language, thus one should be cautions in drawing conclusions on this exercise.

I consider Tables A.14 to A.17 first. Overall my findings suggest that for Mexicans the trade-off induced by an early arrival in the U.S. is milder and that for this group the negative effect of age at arrival on cognitive skills and education is larger. In Table A.14, it is possible to observe that the effect of an extra year in the home country is about one third bigger in absolute value for Mexicans (-0.12+(-0.08) vs -0.12). A similar conclusion emerges from Table A.15 and Table A.16, on English and mathematics knowledge. For example, for paragraph comprehension, an extra year in age at arrival in the U.S. reduces outcomes of Mexicans by 0.12 to 0.13 standard deviations, while the reduction for immigrants from other countries is 0.08 to 0.12. As for illicit behaviors, age at arrival has a larger role for individuals arrived from other countries, while smaller for Mexicans. Thus, for them the trade-off induced by age at arrival is milder.

In Tables A.18 to A.21 I study, instead, the heterogeneity based on the language spoken in the origin country of immigrants. It is important to notice that, given the small sample of individuals coming from countries where English is an official language, the estimates are very noisy and they might be misleading. Particularly when using the empirical model with categorical variables for age at arrival, it is possible, indeed, that the estimates are induced by very few individuals. Furthermore, because of cultural and institutional differences across countries, this exercise should not be considered as the solution for disentangling the linguistic component of the age at arrival effect. First, the effect of age at arrival on schooling, in Table A.18, seems stronger in absolute terms for individuals that speak English in their origin countries. In terms of the effect of age at arrival on linguistic achievements, perhaps surprisingly, there is no big difference between the two groups. For immigrants that speak English in their origin countries, instead, arriving later in the U.S. has no or very limited role in affecting their mathematical scores. To conclude, in Table A.21 I find no differences between the two groups in terms of the effect of age at arrival on social behaviors.

1.6 Robustness checks

In this subsection I check the robustness of the main results performing several tests. First, I check if the presence of both parents in the households affects the estimates. Second, I exclude from the analysis on schooling outcomes immigrant children arrived in the U.S. as dropouts. Third, I control for time trend in the

American educational system. Forth, I run the regressions using a subsample of individuals arrived at the age of 13 or younger.

Presence of both parents in households

The presence of both parents in households is likely correlated with both the development of personality traits of young individuals, as well as with age at arrival. In particular, within families, children arrived when they are younger are also those more likely to suffer the absence of one parent (or both) for a longer period of time. This may be due, for example, to parental separation or death.

To address this concern, in Tables A.1 to A.7 I repeat the analysis controlling for the presence of both parents in the household. Both the magnitude and the significance of the estimates of this exercise are almost identical to those described in the main section.

Excluding immigrants arriving as dropouts

The results on schooling, reported in Table 1.3, might mix two effects. On the one hand, age at arrival might increase the likelihood of leaving school after arriving in the U.S., on the other hand it might reduce the likelihood of re-enrolling after arriving in the U.S. as a dropout.

To possibly separate these two effects, I exclude from the analysis the immigrant children that do not attend U.S. schools. The results are reported in Table A.8. The estimates using the model where age at arrival enters linearly in the regression are marginally smaller than the ones using the whole sample. Furthermore, using the more flexible model, with age at arrival expressed in terms of a categorical variable, the coefficient on individuals arrived between the ages 15 to 17 reduces by about 0.5 years in absolute terms. This indicates that the large difference in educational achievements between immigrants arrived before the age of 6 and those arrived between the ages 15 and 17 are partly driven by the high likelihood of not re-enrolling in school after becoming dropouts in the home countries. Nonetheless, for individuals attending at least one year of formal education in the U.S., an arrival around the age of 16 still accounts for a reduction of 1.15 to 1.8 years of completed education, compared to individuals arrived before the age of 6.

Controlling for time trends in the education system

A possibility for the results on schooling, reported in Table 1.3, is that the age at arrival variable captures time trends in the American educational systems. For example, individuals that arrived younger in the U.S. attended school in more recent years. Thus, if the American educational system evolved in the period under consideration, my estimates could be biased even when considering the within family variation.

First, given that the individuals used in my study are all born within a window of 7 years, it is quite unlikely that major changes in the American schooling system differentially affected their educational outcomes. However, I additionally re-estimate my model using a pooled sample of immigrants and natives, controlling for a time trend. As long as changes in the schooling systems have the same impact on immigrants and natives, this empirical strategy addresses the concern. In particular, the age at arrival effect is estimated via the within immigrant family variation in age at arrival in the U.S., while the time trends in the American educational system is controlled for using the within family variation in year of birth across siblings in native families. I report the result of this robustness test in Table A.9. I conclude from this Table that the main result on schooling is robust to time trends in the American schooling system[12].

[12] This empirical strategy, however, is not useful if improvements in the American educational system impact in a different way immigrants and natives.

Excluding immigrants arrived after the age of 14

One concern of my empirical setting is that children of immigrants can affect the parental decision on the timing of migration. Children in their teenage years, in particular, might be those more likely to influence their parents.

To address this concern, I exclude from my analysis individuals arrived after the age of 14. I only report, in Tables A.10 to A.13, results for the linear model. The coefficients of the more flexible model, with age at arrival defined as categorical variable, are indeed unaffected by this sampling strategy. The magnitudes of the estimated coefficients are very similar to those described in the main section. In general, the coefficients are a bit larger in absolute value, although reducing the sample size increases the standard errors, too. Therefore, most of the coefficients, in these cases, turn out to be insignificant.

1.7 Conclusion

In this paper I studied the effect of age at arrival on the development of children of immigrants' personalities. As the number of immigrants is rising in most of the richest economies, they are becoming increasingly important in shaping the human capital of the average individual in these economies. Young immigrants are a peculiar category in that they do not choose to become immigrants, but they might face more severe challenges and bring additional benefits to their host countries. Indeed, relative to their parents, they need to switch to a new educational system and, on average, they stay in the host countries for longer.

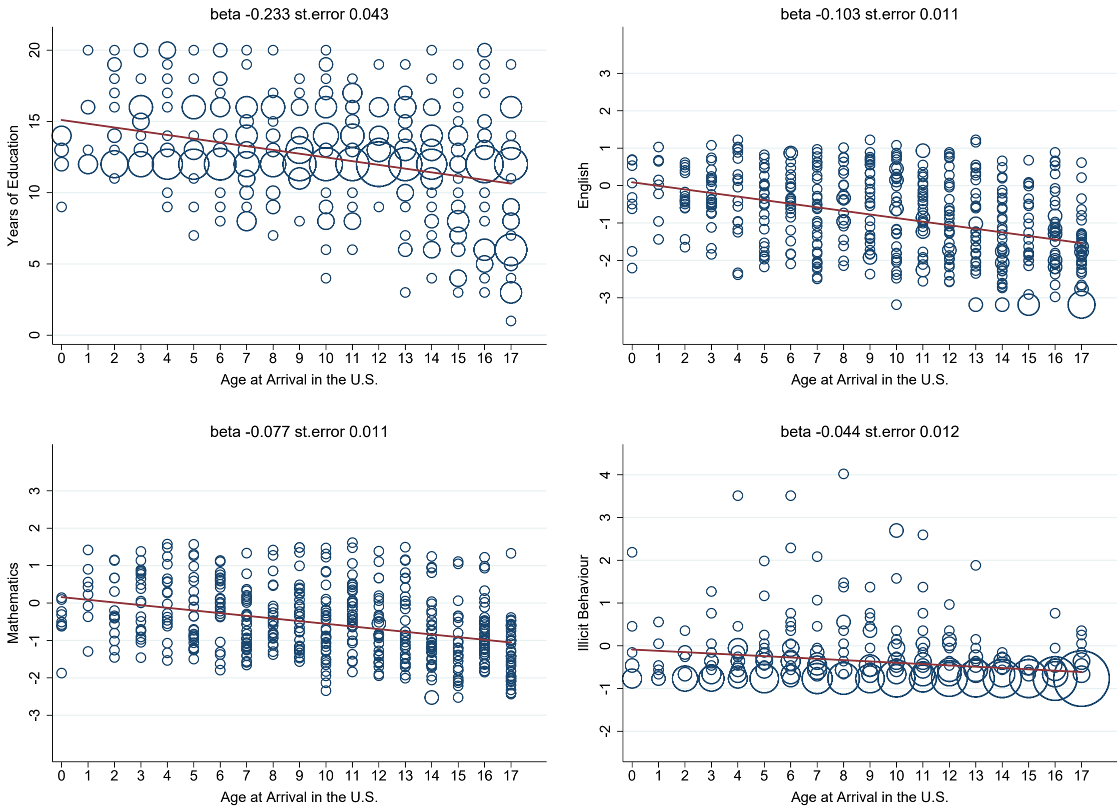
Compared to other studies, I used several measures to capture the multifaceted aspect of human capital. I observed immigrants' cognitive development through their educational achievements and their performance in standardized tests. Moreover, I also studied their attitude in being involved in problematic and anti-social behaviours. Using the variation in age at arrival across siblings living in the same household, I showed that age at arrival has a negative effect on schooling and cognitive outcomes, but it also reduces the likelihood of being involved in illicit activities. In other words, the earlier an immigrant enters in the U.S., the more similar he becomes to the average American, along several human capital dimensions often used in the literature. This suggests the presence of a policy trade-off in targeting immigrant families arriving with their children, as, perhaps contrary to a popular belief, Americans in the NLSY79 tend to behave more illicitly than first generation immigrants. I also showed through indirect estimates that the net effect of age at arrival on labour market outcomes is negative.

As for cognitive skills, the effect is stronger for linguistic ability, and milder for mathematics. I also documented that linguistic ability might be a relevant channel through which age at arrival affects the cognitive achievements, but it is unlikely to be an important one for the effect of age at arrival on illicit activities. From a policy perspective, providing linguistic support to immigrants might then be beneficial to reduce their barriers in the accumulation of cognitive skills, retaining some positive aspects of being a foreign-born individual.

I concluded my analysis looking at the parental selection in age at arrival. In both the NLSY79 and Census data, more educated parents are those lowering the age at arrival of their children. In light of my estimates for the effect of age at arrival on cognitive skills accumulation, this finding has important implications for policies aiming at reducing the achievement gaps between children of poor and rich families. Children raising in poor families might find the support for enhancing their educational attainments particularly beneficial when born in immigrant households.

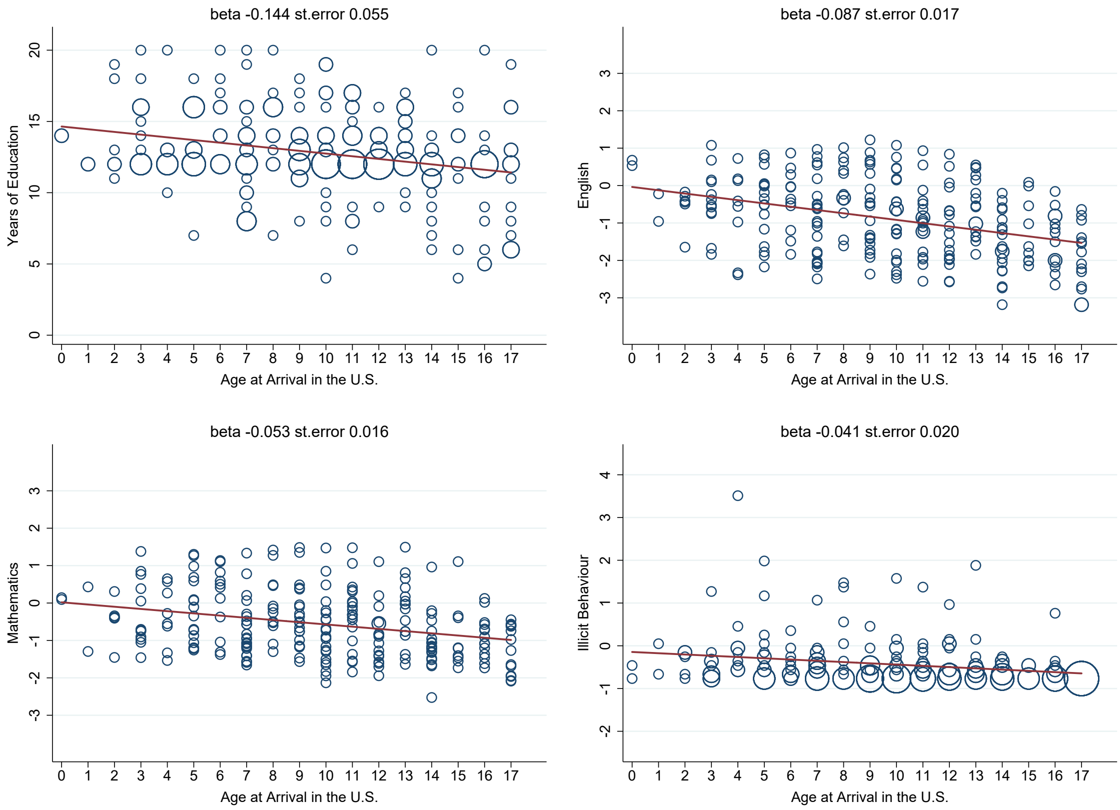
Tables and figures

Figure 1-1: The correlation between different outcomes and age at arrival in the U.S.



These Figures depict the correlation between different immigrants' achievements, measured on the Y-axis, and their age at arrival in the U.S., measured on the X-axis. At the top of each panel, the coefficients and the standard errors of the linear regressions between the two variables are shown. The dimension of the blue circles depends on the frequency of individuals with the same outcome - age at arrival profile. The sample consists of the immigrant population in the NLSY79 arrived in the U.S. at the age of 17 or earlier. The linguistic measure is the sum of the achievements in the word knowledge and paragraph comprehension sections of the ASVAB test, while the mathematical measure is the sum of the achievements in the arithmetic reasoning, numerical operation and high-school mathematics sections of the ASVAB. Each variable, but the educational achievement, has been standardized to have a mean of 0 and a standard deviation of 1 in the NLSY79 population.

Figure 1-2: The correlation between different outcomes and age at arrival in the U.S.



These Figures depict the correlations between different immigrants' achievements, measured on the Y-axis, and their age at arrival in the U.S., measured on the X-axis. At the top of each panel, the coefficients and the standard errors of the linear regressions between the two variables are shown. The dimension of the blue circles depends on the frequency of individuals with the same outcome - age at arrival profile. The sample consists of the immigrant population in the NLSY79 arrived in the U.S. at the age of 17 or earlier, with at least one brother or sister in the sample. The linguistic measure is the sum of the achievements in the word knowledge and paragraph comprehension sections of the ASVAB test, while the mathematical measure is the sum of the achievements in the arithmetic reasoning, numerical operation and high-school mathematics sections of the ASVAB. Each variable, but the educational achievement, has been standardized to have a mean of 0 and a standard deviation of 1 in the NLSY79 population.

Table 1.1: Summary statistics; NLSY79

Sample	Native	1st Generation	
		Whole	Siblings
Years of Schooling	13.50	13.06	13.05
Word Knowledge	-0.02	-0.75	-0.79
Paragraph Comprehension	-0.01	-0.70	-0.70
Arithmetic Reasoning	-0.02	-0.48	-0.45
Numerical Operations	-0.02	-0.34	-0.32
Mathematics Knowledge	-0.03	-0.38	-0.32
Illicit Behaviours	0.00	-0.35	-0.40
Age At Arrival		9.90	9.71
Individuals	10059	537	256

Table 1.2: The correlation between educational achievement and age at arrival; NLSY79 and Census

Dataset	NLSY79		Census			
Sample	Whole	Siblings	All Years, Whole	All Years, NLSY Years of Birth	2000, Whole	2000, NLSY Years of Birth
	(1)	(2)	(3)	(4)		
Age At Arrival	-0.26***	-0.14***	-0.19***	-0.21***	-0.21***	-0.23***
	(0.03)	(0.06)	(0.00)	(0.00)	(0.00)	(0.00)
N	537	256	614800	141314	304626	70205

The Table above reports the results of regressions of years of education on immigrants' age at arrival in the U.S.. As for NLSY data, Whole refers to immigrant individuals arrived in the U.S. at the age of 17 or younger, while Siblings refers to immigrant individuals with at least one brother or sister in the same sample. As for Census data, Whole refers to immigrant individuals in prime age, arrived in the U.S. at the age of 17 or younger. NLSY Years of Birth refers to immigrant individuals in prime age arrived in the U.S. at the age of 17 or younger, and belonging to the same cohorts of immigrants in the NLSY79. *, ** , and *** denote significance at the 10%, 5%, 1% level, respectively.

Table 1.3: The effect of age at arrival on years of education; ordinary least squares and family fixed effects regressions

Sample	Full						Same Year of Arrival					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Age At Arrival	-0.14**		-0.14**		-0.15		-0.17**		-0.16**		-0.20	
	(0.06)		(0.06)		(0.09)		(0.08)		(0.08)		(0.13)	
Age At Arrival 6-8		0.60		0.12		0.12		0.57		-0.31		-0.37
		(0.91)		(0.45)		(0.50)		(1.05)		(0.47)		(0.52)
Age At Arrival 9-11		-0.22		-0.92*		-0.91*		-0.24		-1.41**		-1.51**
		(0.61)		(0.49)		(0.54)		(0.78)		(0.63)		(0.72)
Age At Arrival 12-14		-0.77		-1.33**		-1.31*		-0.99		-1.76**		-1.97**
		(0.57)		(0.58)		(0.78)		(0.68)		(0.71)		(0.97)
Age At Arrival 15-17		-2.27**		-1.69**		-1.66*		-2.66**		-2.09**		-2.37**
		(0.88)		(0.72)		(0.98)		(1.01)		(0.82)		(1.18)
Family F.E.	No	No	Yes	Yes	Yes	Yes	No	No	Yes	Yes	Yes	Yes
Birth Order	No	No	No	No	Yes	Yes	No	No	No	No	Yes	Yes
N	256	256	256	256	256	256	213	213	213	213	213	213

The Table above reports the results of regressions of years of education on immigrants' age at arrival in the U.S.. In odd columns I report the results of a linear regression model, whereas even columns display the results of a model using categorical variables for intervals of age at arrival in the U.S.. In each specification I control for a sex indicator and, when specified, for family fixed effects and a birth order control. In the first six columns the sample consists of immigrant individuals arrived in the U.S. at the age of 17 or younger, with at least one sibling in the sample. In columns seven to twelve the sample consists of immigrant individuals arrived in the U.S. at the age of 17 or younger, with at least one sibling in the sample, with the additional requirement that each sibling arrived in the U.S. in the same calendar year. Heteroskedasticity robust standard errors are clustered at the family level and they are shown in parenthesis. *, ** , and *** denote significance at the 10%, 5%, 1% level, respectively.

Table 1.4: The effect of age at arrival on word knowledge; ordinary least squares and family fixed effects regressions

Sample	Full						Same Year of Arrival					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Age At Arrival	-0.06**		-0.06**		-0.05*		-0.05*		-0.05*		-0.04	
	(0.02)		(0.02)		(0.03)		(0.02)		(0.03)		(0.03)	
Age At Arrival 6-8		-0.03		-0.03		-0.01		-0.06		-0.06		-0.06
		(0.14)		(0.14)		(0.15)		(0.15)		(0.13)		(0.15)
Age At Arrival 9-11		-0.17		-0.17		-0.14		-0.15		-0.15		-0.12
		(0.12)		(0.11)		(0.14)		(0.12)		(0.10)		(0.16)
Age At Arrival 12-14		-0.50***		-0.50***		-0.47*		-0.48**		-0.48***		-0.49**
		(0.19)		(0.19)		(0.25)		(0.19)		(0.18)		(0.25)
Age At Arrival 15-17		-0.60**		-0.60**		-0.53		-0.51**		-0.51**		-0.51
		(0.25)		(0.25)		(0.33)		(0.25)		(0.25)		(0.35)
Family F.E.	No	No	Yes	Yes	Yes	Yes	No	No	Yes	Yes	Yes	Yes
Birth Order	No	No	No	No	Yes	Yes	No	No	No	No	Yes	Yes
N	4511	4511	4511	4511	4511	4511	4476	4476	4476	4476	4476	4476

The Table above reports the results of regressions of word knowledge on immigrants' age at arrival in the U.S.. In odd columns I report the results of a linear regression model, whereas even columns display the results of a model using categorical variables for intervals of age at arrival in the U.S.. In each specification I control for a sex indicator and, when specified, for family fixed effects and a birth order control. All these controls are interacted with a dummy variable indicating the immigration status. The sample in each column consists of native and immigrant siblings. In the immigrant sample of the first six columns I keep individuals arrived in the U.S. at the age of 17 or younger, with at least one sibling in the sample. In columns seven to twelve the immigrant sample consists of immigrant individuals arrived in the U.S. at the age of 17 or younger, with at least one sibling in the sample, with the additional requirement that each sibling arrived in the U.S. in the same calendar year. Heteroskedasticity robust standard errors are clustered at the family level and they are shown in parenthesis. *, ** , and *** denote significance at the 10%, 5%, 1% level, respectively.

Table 1.5: The effect of age at arrival on paragraph comprehension; ordinary least squares and family fixed effects regressions

Sample	Full						Same Year of Arrival					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Age At Arrival	-0.07**		-0.07**		-0.09**		-0.07**		-0.07**		-0.12**	
	(0.03)		(0.03)		(0.04)		(0.03)		(0.03)		(0.06)	
Age At Arrival 6-8		0.15		0.15		0.05		0.08		0.08		-0.06
		(0.16)		(0.15)		(0.17)		(0.18)		(0.16)		(0.18)
Age At Arrival 9-11		-0.17		-0.17		-0.29		-0.17		-0.17		-0.37
		(0.17)		(0.17)		(0.19)		(0.22)		(0.21)		(0.25)
Age At Arrival 12-14		-0.67***		-0.67***		-1.00***		-0.68**		-0.68**		-1.15***
		(0.23)		(0.23)		(0.28)		(0.27)		(0.26)		(0.34)
Age At Arrival 15-17		-0.80***		-0.80***		-1.23***		-0.85***		-0.85***		-1.45***
		(0.29)		(0.29)		(0.37)		(0.32)		(0.32)		(0.42)
Family F.E.	No	No	Yes	Yes	Yes	Yes	No	No	Yes	Yes	Yes	Yes
Birth Order	No	No	No	No	Yes	Yes	No	No	No	No	Yes	Yes
N	4511	4511	4511	4511	4511	4511	4476	4476	4476	4476	4476	4476

The Table above reports the results of regressions of paragraph comprehension on immigrants' age at arrival in the U.S.. In odd columns I report the results of a linear regression model, whereas even columns display the results of a model using categorical variables for intervals of age at arrival in the U.S.. In each specification I control for a sex indicator and, when specified, for family fixed effects and a birth order control. All these controls are interacted with a dummy variable indicating the immigration status. The sample in each column consists of native and immigrant siblings. In the immigrant sample of the first six columns I keep individuals arrived in the U.S. at the age of 17 or younger, with at least one sibling in the sample. In columns seven to twelve the immigrant sample consists of immigrant individuals arrived in the U.S. at the age of 17 or younger, with at least one sibling in the sample, with the additional requirement that each sibling arrived in the U.S. in the same calendar year. Heteroskedasticity robust standard errors are clustered at the family level and they are shown in parenthesis. *, ** , and *** denote significance at the 10%, 5%, 1% level, respectively.

Table 1.6: The effect of age at arrival on arithmetic reasoning; ordinary least squares and family fixed effects regressions

Sample	Full						Same Year of Arrival					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Age At Arrival	-0.02		-0.02		-0.05		-0.01		-0.01		-0.06	
	(0.03)		(0.03)		(0.03)		(0.04)		(0.04)		(0.05)	
Age At Arrival 6-8		-0.10		-0.10		-0.19		-0.04		-0.04		-0.16
		(0.14)		(0.12)		(0.15)		(0.16)		(0.13)		(0.18)
Age At Arrival 9-11		-0.11		-0.11		-0.22		-0.03		-0.03		-0.21
		(0.19)		(0.16)		(0.22)		(0.24)		(0.20)		(0.29)
Age At Arrival 12-14		-0.21		-0.21		-0.53**		-0.14		-0.14		-0.58*
		(0.29)		(0.25)		(0.26)		(0.32)		(0.27)		(0.33)
Age At Arrival 15-17		-0.19		-0.19		-0.60		-0.15		-0.15		-0.70
		(0.31)		(0.29)		(0.40)		(0.35)		(0.31)		(0.49)
Family F.E.	No	No	Yes	Yes	Yes	Yes	No	No	Yes	Yes	Yes	Yes
Birth Order	No	No	No	No	Yes	Yes	No	No	No	No	Yes	Yes
N	4511	4511	4511	4511	4511	4511	4476	4476	4476	4476	4476	4476

The Table above reports the results of regressions of arithmetic reasoning on immigrants' age at arrival in the U.S.. In odd columns I report the results of a linear regression model, whereas even columns display the results of a model using categorical variables for intervals of age at arrival in the U.S.. In each specification I control for a sex indicator and, when specified, for family fixed effects and a birth order control. All these controls are interacted with a dummy variable indicating the immigration status. The sample in each column consists of native and immigrant siblings. In the immigrant sample of the first six columns I keep individuals arrived in the U.S. at the age of 17 or younger, with at least one sibling in the sample. In columns seven to twelve the immigrant sample consists of immigrant individuals arrived in the U.S. at the age of 17 or younger, with at least one sibling in the sample, with the additional requirement that each sibling arrived in the U.S. in the same calendar year. Heteroskedasticity robust standard errors are clustered at the family level and they are shown in parenthesis. *, ** , and *** denote significance at the 10%, 5%, 1% level, respectively.

Table 1.7: The effect of age at arrival on numerical operations; ordinary seast squares and family fixed effects regressions

Sample	Full						Same Year of Arrival					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Age At Arrival	-0.04		-0.04		-0.04		-0.05		-0.05		-0.06	
	(0.04)		(0.04)		(0.05)		(0.04)		(0.04)		(0.06)	
Age At Arrival 6-8		-0.03		-0.03		-0.04		-0.16		-0.16		-0.19
		(0.28)		(0.27)		(0.29)		(0.30)		(0.29)		(0.31)
Age At Arrival 9-11		-0.06		-0.06		-0.07		-0.21		-0.21		-0.26
		(0.30)		(0.31)		(0.32)		(0.35)		(0.36)		(0.40)
Age At Arrival 12-14		-0.50		-0.50		-0.54		-0.70*		-0.70*		-0.83
		(0.34)		(0.36)		(0.42)		(0.39)		(0.42)		(0.52)
Age At Arrival 15-17		-0.37		-0.37		-0.41		-0.44		-0.44		-0.60
		(0.50)		(0.54)		(0.62)		(0.54)		(0.58)		(0.70)
Family F.E.	No	No	Yes	Yes	Yes	Yes	No	No	Yes	Yes	Yes	Yes
Birth Order	No	No	No	No	Yes	Yes	No	No	No	No	Yes	Yes
N	4511	4511	4511	4511	4511	4511	4476	4476	4476	4476	4476	4476

The Table above reports the results of regressions of numerical operations ability on immigrants' age at arrival in the U.S.. In odd columns I report the results of a linear regression model, whereas even columns display the results of a model using categorical variables for intervals of age at arrival in the U.S.. In each specification I control for a sex indicator and, when specified, for family fixed effects and a birth order control. All these controls are interacted with a dummy variable indicating the immigration status. The sample in each column consists of native and immigrant siblings. In the immigrant sample of the first six columns I keep individuals arrived in the U.S. at the age of 17 or younger, with at least one sibling in the sample. In columns seven to twelve the immigrant sample consists of immigrant individuals arrived in the U.S. at the age of 17 or younger, with at least one sibling in the sample, with the additional requirement that each sibling arrived in the U.S. in the same calendar year. Heteroskedasticity robust standard errors are clustered at the family level and they are shown in parenthesis. *, ** , and *** denote significance at the 10%, 5%, 1% level, respectively.

Table 1.8: The effect of age at arrival on mathematics knowledge; ordinary least squares and family fixed effects regressions

Sample	Full						Same Year of Arrival					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Age At Arrival	-0.04		-0.04		-0.03		-0.03		-0.03		-0.00	
	(0.02)		(0.02)		(0.04)		(0.03)		(0.03)		(0.04)	
Age At Arrival 6-8		-0.24		-0.24		-0.24		-0.27		-0.27		-0.25
		(0.17)		(0.17)		(0.18)		(0.19)		(0.19)		(0.20)
Age At Arrival 9-11		-0.27		-0.27		-0.25		-0.19		-0.19		-0.14
		(0.18)		(0.19)		(0.21)		(0.23)		(0.24)		(0.27)
Age At Arrival 12-14		-0.48**		-0.48**		-0.49		-0.42		-0.42		-0.39
		(0.22)		(0.23)		(0.30)		(0.26)		(0.27)		(0.35)
Age At Arrival 15-17		-0.40		-0.40		-0.38		-0.34		-0.34		-0.27
		(0.26)		(0.26)		(0.36)		(0.29)		(0.30)		(0.42)
Family F.E.	No	No	Yes	Yes	Yes	Yes	No	No	Yes	Yes	Yes	Yes
Birth Order	No	No	No	No	Yes	Yes	No	No	No	No	Yes	Yes
N	4511	4511	4511	4511	4511	4511	4476	4476	4476	4476	4476	4476

The Table above reports the results of regressions of mathematics knowledge on immigrants' age at arrival in the U.S.. In odd columns I report the results of a linear regression model, whereas even columns display the results of a model using categorical variables for intervals of age at arrival in the U.S.. In each specification I control for a sex indicator and, when specified, for family fixed effects and a birth order control. All these controls are interacted with a dummy variable indicating the immigration status. The sample in each column consists of native and immigrant siblings. In the immigrant sample of the first six columns I keep individuals arrived in the U.S. at the age of 17 or younger, with at least one sibling in the sample. In columns seven to twelve the immigrant sample consists of immigrant individuals arrived in the U.S. at the age of 17 or younger, with at least one sibling in the sample, with the additional requirement that each sibling arrived in the U.S. in the same calendar year. Heteroskedasticity robust standard errors are clustered at the family level and they are shown in parenthesis. *, ** , and *** denote significance at the 10%, 5%, 1% level, respectively.

Table 1.9: The effect of age at arrival on illicit behaviours; ordinary least squares and family fixed effects regressions

Sample	Full						Same Year of Arrival					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Age At Arrival	-0.05***		-0.06		-0.09		-0.05***		-0.08		-0.15	
	(0.01)		(0.06)		(0.08)		(0.01)		(0.07)		(0.11)	
Age At Arrival 6-8		-0.34**		-0.49		-0.54		-0.32		-0.55		-0.66
		(0.17)		(0.37)		(0.37)		(0.20)		(0.47)		(0.48)
Age At Arrival 9-11		-0.36***		-0.73		-0.83*		-0.32**		-0.97		-1.22*
		(0.13)		(0.48)		(0.50)		(0.16)		(0.62)		(0.66)
Age At Arrival 12-14		-0.54***		-0.55		-0.75		-0.58***		-0.89		-1.32**
		(0.14)		(0.46)		(0.50)		(0.15)		(0.56)		(0.64)
Age At Arrival 15-17		-0.71***		-0.65		-0.93*		-0.71***		-0.98*		-1.58**
		(0.13)		(0.47)		(0.54)		(0.15)		(0.57)		(0.70)
Family F.E.	No	No	Yes	Yes	Yes	Yes	No	No	Yes	Yes	Yes	Yes
Birth Order	No	No	No	No	Yes	Yes	No	No	No	No	Yes	Yes
N	4158	4158	4158	4158	4158	4158	4126	4126	4126	4126	4126	4126

The Table above reports the results of regressions of illicit behaviours on immigrants' age at arrival in the U.S.. In odd columns I report the results of a linear regression model, whereas even columns display the results of a model using categorical variables for intervals of age at arrival in the U.S.. In each specification I control for a sex indicator and, when specified, for family fixed effects and a birth order control. All these controls are interacted with a dummy variable indicating the immigration status. The sample in each column consists of native and immigrant siblings. In the immigrant sample of the first six columns I keep individuals arrived in the U.S. at the age of 17 or younger, with at least one sibling in the sample. In columns seven to twelve the immigrant sample consists of immigrant individuals arrived in the U.S. at the age of 17 or younger, with at least one sibling in the sample, with the additional requirement that each sibling arrived in the U.S. in the same calendar year. Heteroskedasticity robust standard errors are clustered at the family level and they are shown in parenthesis. *, ** , and *** denote significance at the 10%, 5%, 1% level, respectively.

Table 1.10: The effect of age at arrival on several outcomes controlling for english ability; whole sample

	Schooling		Arithmetic		Num. Operations		Maths Knowledge		Illicit	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Age At Arrival	-0.11		-0.01		-0.00		0.01		-0.10	
	(0.08)		(0.03)		(0.05)		(0.03)		(0.09)	
Age At Arrival 6-8		0.03		-0.20		-0.05		-0.24		-0.54
		(0.49)		(0.15)		(0.28)		(0.17)		(0.38)
Age At Arrival 9-11		-0.79		-0.09		0.05		-0.12		-0.89*
		(0.54)		(0.22)		(0.33)		(0.21)		(0.52)
Age At Arrival 12-14		-0.74		-0.08		-0.14		-0.07		-0.88*
		(0.79)		(0.24)		(0.42)		(0.30)		(0.50)
Age At Arrival 15-17		-1.07		-0.06		0.07		0.12		-1.06*
		(0.94)		(0.37)		(0.61)		(0.35)		(0.55)
Word Know.	0.41*	0.43*	0.34***	0.34***	0.28***	0.28***	0.39***	0.39***	-0.06	-0.06
	(0.23)	(0.25)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)
Parag. Comp.	0.58***	0.58***	0.29***	0.29***	0.27***	0.27***	0.24***	0.24***	0.06*	0.06*
	(0.21)	(0.22)	(0.02)	(0.02)	(0.03)	(0.03)	(0.02)	(0.02)	(0.04)	(0.04)
Family F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Birth Order	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	236	236	4511	4511	4511	4511	4511	4511	4028	4028

The Table above reports the results of regressions of several outcomes on immigrants' age at arrival in the U.S.. In odd columns I report the results of a linear regression model, whereas even columns display the results of a model using categorical variables for intervals of age at arrival in the U.S.. In each specification I control for a sex indicator, family fixed effects and a birth order control. All these controls are interacted with a dummy variable indicating the immigration status. The sample in the first two columns consists of immigrant siblings, while the sample in each other column consists of native and immigrant siblings. In the immigrant sample I keep individuals arrived in the U.S. at the age of 17 or younger, with at least one sibling in the sample. Heteroskedasticity robust standard errors are clustered at the family level and they are shown in parenthesis. *, ** , and *** denote significance at the 10%, 5%, 1% level, respectively.

Table 1.11: The effect of age at arrival on several outcomes controlling for english ability; same year of arrival sample

	Schooling		Arithmetic		Num. Operations		Maths Knowledge		Illicit	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Age At Arrival	-0.18		-0.01		-0.02		0.04		-0.15	
	(0.11)		(0.04)		(0.06)		(0.04)		(0.11)	
Age At Arrival 6-8		-0.43		-0.12		-0.16		-0.22		-0.68
		(0.52)		(0.17)		(0.32)		(0.20)		(0.48)
Age At Arrival 9-11		-1.38**		-0.06		-0.13		0.00		-1.24*
		(0.65)		(0.30)		(0.43)		(0.28)		(0.67)
Age At Arrival 12-14		-1.36		-0.08		-0.39		0.08		-1.30**
		(0.91)		(0.31)		(0.55)		(0.37)		(0.65)
Age At Arrival 15-17		-1.73		-0.11		-0.07		0.27		-1.57**
		(1.09)		(0.46)		(0.72)		(0.43)		(0.71)
Word Know.	0.43	0.44	0.34***	0.34***	0.28***	0.28***	0.39***	0.39***	-0.06	-0.06
	(0.29)	(0.32)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)
Parag. Comp.	0.56**	0.59**	0.29***	0.29***	0.27***	0.27***	0.24***	0.24***	0.06*	0.06*
	(0.23)	(0.24)	(0.02)	(0.02)	(0.03)	(0.03)	(0.02)	(0.02)	(0.04)	(0.04)
Family F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Birth Order	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	201	201	4476	4476	4476	4476	4476	4476	3998	3998

The Table above reports the results of regressions of several outcomes on immigrants' age at arrival in the U.S.. In odd columns I report the results of a linear regression model, whereas even columns display the results of a model using categorical variables for intervals of age at arrival in the U.S.. In each specification I control for a sex indicator, family fixed effects and a birth order control. All these controls are interacted with a dummy variable indicating the immigration status. The sample in the first two columns consists of immigrant siblings, while the sample in each other column consists of native and immigrant siblings. In the immigrant sample I keep individuals arrived in the U.S. in the same calendar year, at the age of 17 or younger, with at least one sibling in the sample. Heteroskedasticity robust standard errors are clustered at the family level and they are shown in parenthesis. *, ** , and *** denote significance at the 10%, 5%, 1% level, respectively.

Table 1.12: The parental selection in age at arrival

Dataset	NLSY79				Census			
Sample	Whole	Siblings			All Years		2000	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mother's Education	-0.14	-0.13	-0.15	-0.03	-0.11***	-0.12***	-0.12***	-0.13***
	(0.11)	(0.16)	(0.16)	(0.20)	(0.01)	(0.01)	(0.01)	(0.01)
Father's Education	-0.03	-0.17	-0.18	-0.13	-0.06***	-0.07***	-0.07***	-0.08***
	(0.12)	(0.16)	(0.16)	(0.21)	(0.01)	(0.01)	(0.01)	(0.01)
Family Income			0.00	0.00				
			(0.00)	(0.00)				
Cohort of Arrival F.E.	No	No	No	No	No	Yes	No	Yes
N	537	256	256	148	335202	335202	142921	142921

The Table above reports the results of regressions of immigrants' age at arrival on parental education. As for NLSY data, Whole refers to immigrant individuals arrived in the U.S. at the age of 17 or younger, while Siblings refers to immigrant individuals arrived in the U.S. at the age of 17 or younger with at least one brother or sister in the same sample. *, ** , and *** denote significance at the 10%, 5%, 1% level, respectively.

Chapter 2

The Immigrant American Dream

2.1 Introduction

The concept of the *American Dream* – as described by James Truslow Adams (1931) – refers to a country where life should be better and richer and fuller for everyone, with opportunity for each according to ability or achievement regardless of social class or circumstances of birth. According to Truslow Adams (1931): "the economic motive was unquestionably powerful, often dominant, in the minds of those who took part in the great migration, but mixed with this was also frequently present the hope of a better and freer life, a life in which a man might think as he would and develop as he willed." It is these motives and the incentives that are associated with the extremely risky decision to pack everything up and move to a different country where everything, often including the language, is new that have been identified as the key factor at the origin of the U.S. social and economic success. Although potentially highly controversial, there is no reason whatsoever to doubt that similar motives are at the foundation of the most recent flows of immigration to the U.S.. There is also no reason to doubt that results would be similar and that all flaws of migration would contribute and boost the U.S. social and economic success.

In this paper we ask whether the immigrant American dream is fulfilled and how many generations it takes to achieve it. We also ask what this means for the

American society and its economy in terms of boosting its overall level of education, economic success and social spirit. Using a random sample of the American youth population in the 1980s, we show that first generation immigrants suffer educational and cognitive skill gaps, but behave less illicitly than the local population. Second and, even more so, third generation immigrants outperform the local population in terms of educational and cognitive achievements. We also document that immigrants tend to have less children than Americans. Furthermore, following the different immigrant cohorts over time, we show that it takes two generations to complete the educational and cognitive immigrant assimilation, while it might take longer for other social skills.

For this study we use the NLSY79, where we can observe a broad set of measures of cognitive and social skills for a random sample of individuals in the U.S., who were between the age of 14 and 21 in 1979. For these individuals we know whether themselves, their parents or their paternal grandfather were born in the U.S.. This allows us to identify them as first, second or third generation immigrants. By comparing their accomplishments to the ones of individuals who were born in the U.S. (and their parents and grand-parents were born in the U.S.), we can show whether the first, second or third generation immigrants under or over perform relative to the U.S. population.

The cross-sectional intergenerational dynamics does not, however, provide a reliable estimate of the true intergenerational immigrant assimilation. First, different immigrant cohorts might be characterized by different potential outcomes, or by different assimilation trajectories. Second, even if the selection into migration is constant over time, the selection of return migrants is not necessarily random. Using the NLSY79, we can address both concerns. In the NLSY79, indeed, we also observe the educational, cognitive and social outcomes of children of female individuals. That is, for each immigrant cohort, we effectively observe two immigrant generations. Taking the difference between children and mother's outcomes, we thus measure to what extent the intergenerational human capital dynamics of immigrants differs relative to the intergenerational dynamics of natives. Our findings suggest that the process of immigrant educational and cognitive assimilation exhausts in two generations. For other social outcomes, for example the attitude towards pregnancy, it takes longer.

We also check the heterogeneity of our results based on first generation immigrant mothers' age at arrival and on the origin country of parents of second generation immigrant mothers. We show that the earlier a first generation immigrant mother enters the U.S., the more similar the intergenerational dynamics is to the dynamics of native child-mother pairs. Furthermore, for second generation immigrant mothers, when both maternal parents are immigrants, the child-mother intergenerational dynamics looks more similar to the dynamics of child-mother pairs with the mother being first generation immigrant. As the majority of immigrants in the U.S. are from Mexico, we also check that Mexican families drive only part of our results.

Using principal component analysis we provide suggestive evidence that the immigrant catch up along several human capital dimensions might be driven by a multifaceted and complex phenomenon. Acquiring proficiency in the local language might, for example, be important for immigrant assimilation in the U.S., but it is unlikely to be the only relevant factor. Moreover, our estimates of the U.S. intergenerational immigrant assimilation are consistent with deterioration in the selection of immigrants between the late 1800s /early 1900s and 1980s.

Our paper is related to the large economic literature on the integration of immigrants and their offspring in several host countries. Notable examples are Chiswick (1977), Borjas (1993), Trejo (1997), Riphahn (2003), Chiswick and DebBurman (2004), Algan et al. (2010), Dustmann et al. (2012). We add to this literature for three main reasons. First, instead of focusing on the educational and labor market performance of immigrants in the U.S., we use a broader set of pre-labor market outcomes. On the one hand, these outcomes, such as the performance in standardized cognitive tests and the attitude towards illicit behaviours, are often treated as unobservable in the immigration literature. On the other, these measures are less likely to be affected by selection into and discrimination on the labor market. Second, while most studies observe the immigrant outcomes up to the second generation, we are able to track the immigrant performance up to the fourth generation[1]. Third, to the best of our knowledge our study is the first to observe the same immigrant cohort across different immigrant generations. From a methodological perspective, our work is similar to Borjas (1985), although this study focuses on first generation immigrants and cannot tackle return migration.

The rest of the paper is organized as follows. Section 2.2 describes the dataset, the main variables of interest and the sampling strategy. Section 2.3 explains the econometric problem and how to use the NLSY79 to address the issue. Section 2.4 reports the main results. Section 2.5 provides suggestive evidence on the complexity of immigrant assimilation and on the evolution of the U.S. immigrant selection. Section 2.6 checks the robustness of the results. Section 2.7 concludes the paper.

2.2 Data and sample selection

For our analysis we use the National Longitudinal Survey of Youth 1979 (NLSY79) and the NLSY79 Children and Young Adults. The NLSY79 is a representative survey of 12686 individuals living in the U.S., who were 14 to 21-year-old at the time of their first survey, in 1979. The survey has been conducted annually until 1994, and then biannually. Importantly, starting from 1986, information on children of women in the NLSY79 has been collected in the NLSY79 Chilren and Young Adults. In each calendar year the participants to the two surveys have responded to a comprehensive set of questions that offer a picture of their cognitive and non-cognitive development, social attitudes, labour market outcomes and parenting.

In both datasets we collect information on schooling achievements, performance in standardized cognitive tests, problematic behaviours, family structure and parenting. Even if some information on children does not harmonize perfectly with

[1]That is, for children of third generation immigrant mothers in the original NLSY79 dataset.

the one collected for their mothers, the two surveys supposedly convey similar phenomena. For example, to measure the development of cognitive skills, people in the NLSY79 took the Armed Services Vocational Aptitude Battery (ASVAB) test in 1979. The ASVAB is the test administered by the U.S. Military Entrance Processing Command to assess the cognitive skills of people that want to join the army. The 10 sections of the ASVAB test are: general science, arithmetic reasoning, word knowledge, paragraph comprehension, mathematics knowledge, electronics information, automotive and shop information, mechanical comprehension, coding speed, numerical operations. In the NLSY79 Chilren and Young Adults, instead, the The Peabody Individual Achievement Test (PIAT) is used. As explained in the guide to the data[2], the "PIAT is a wide-range measure of academic achievement for children aged five and over. It is among the most widely used brief assessment of academic achievement with high test-retest reliability and concurrent validity". In particular, the NLSY79 Children and Young Adults includes three sections of the full PIAT: the mathematics, reading recognition, and reading comprehension assessments. In our analysis we thus assume that the ASVAB and the PIAT are similarly useful in summarizing the mathematics and English knowledge of children and teenagers.

Our study benefits from the two NLS datasets for several reasons. Both sources provide different measures of cognitive and social assimilation of immigrants that are not available in other datasets, for example the performance in standardized tests or the tendency to undertake problematic, illicit and anti-social behaviours. Individuals are observed since they are children or teenagers. We can thus study the development of early-life skills, rather than focusing on measures that might reflect complex underlying factors, possibly different across immigrant generations. For example, labor market outcomes might reflect selection into and discrimination on the labor market, as well as social assimilation in the local environment. These factors should, however, be less important in the determination of the performance in a mathematics

[2]The link to the guide is: `https://www.nlsinfo.org/content/cohorts/nlsy79-children/topical-guide/assessments/piat-mathematics`.

test taken at the age of 18. The combination of the NLSY79 and the NLSY79 Children and Young Adults datasets is crucial. Linking the outcomes of mothers and their children, we can make inference on the process of immigrants' assimilation, avoiding the problems associated to the use of cross-sectional or repeated cross-sectional data. We discuss the last point in detail in the empirical section of the paper.

We characterize the immigration status of individuals in the two surveys starting from the NLSY79 dataset. In the NLSY79 we use information on individuals' country of birth, his parental countries of birth and his paternal grandfather's country of birth[3]. To be more specific, we define a person as native American if he is born in the U.S. from two parents born in the U.S., and if his paternal grandfather is also born in the U.S.. An individual is first generation immigrant if the country of origin of himself, his parents and his paternal grandfather is not the U.S.. He is second generation immigrant if he is born in the U.S., but either his mother or his father[4] or both are born abroad. He is third generation immigrant if his and his parents' country of origin is the U.S., whereas his paternal grandfather's is not the U.S.. To avoid that certain individual characteristics might be the causes, rather than the consequences of immigration, we drop first generation immigrants arrived in the U.S. after the age of 17. Out of a total of 12686 individuals in the NLSY79, we focus on the 12325 that fall into one of the categories described above. 10059 individuals are classified as natives, 537 as first generation, 756 as second generation and 973 as third generation immigrants. Depending on their mothers' immigration status, we define accordingly the status of their children. In total, out of the original 11521 individuals in the NLSY79 Children and Young Adults dataset we use 10223 individuals. 8341 are children of natives, 522 children of first generation, 635 children of second generation and 725 children of third generation immigrant mothers.

To measure the cognitive development of individuals in the NLSY79 dataset we

[3]We do not observe the country of origin of the other grandparents.

[4]If only the father is born abroad we also require his paternal grandfather to be born abroad.

rely on information on their education and on the results in the English and mathematics sections of the ASVAB test. As for English, we use the combined word knowledge - paragraph comprehension age-adjusted scores provided by the NLS, whereas for mathematics we use the combined arithmetic reasoning - mathematics knowledge age-adjusted scores. In terms of social assimilation, we focus on behavioral problems and some measures of family structure. We construct an illicit activity index, that is a proxy for engaging in illicit activities when surveyed in the 1980. The index is based on 23 questions; to obtain the index, we add the different indicators for being involved in the particular behavior described in the question and we divide by the number of questions. Higher scores are thus associated to more severe behavioral problems. We also use information on the presence of both parents in the household when the respondent was 14 year-old. For female respondents, we collect information on parenting. In particular, we observe the number of children they have, as well as the ages at which they gave birth.

To measure the cognitive development of individuals in the NLSY79 Children and Young Adults dataset we use the age-adjusted results in the mathematics, reading recognition, and reading comprehension sections of the PIAT assessment. To describe their attitudes towards problematic behaviours we use the Behaviour Problem Index (BPI). All other measures, including the ones on schooling outcomes and parenting, are identical to the ones for their mothers in the NLSY79.

2.3 Econometric model

We assume that the outcome of an individual t belonging to a group i is determined by the following equation:

$$Y_{it} = \beta_0 + \sum_{g>0} \gamma_g \mathbb{1}\{G_{it} \geq g\} + \beta X_{it} + \overbrace{\theta_i + \epsilon_{it}}^{u_{it}} \tag{2.1}$$

where G_{it} is the immigrant generation of the individual, X_{it} is a set of factors that can vary at the individual level, and u_{it} represents a shock. The shock term is the sum of a group-specific component, θ_i and an idiosyncratic term ϵ_{it}. The sum $\sum_g \gamma_g \mathbb{1}\{G_{it} \geq g\}$ describes the intergenerational assimilation process of individual t and his ancestors in the U.S.. Conditionally on the term X_{it} and the group-specific component θ_i, β_0 can then be interpreted as the average outcome of individuals born in the U.S. from ancestors that are also U.S. citizens, whereas each γ_g represents the speed of assimilation from one immigrant generation to the next. In other words, we are assuming that the outcome of an immigrant of generation g not only depends on his ability to take advantage of the American resources relative to natives, γ_g, but also on the process of assimilation of his ancestors in the host country, measured by γ_{g-1}, γ_{g-2}, ..., γ_1.

The purpose of our study is to estimate the parameters describing the intergenerational assimilation of immigrants relative to natives, up to the fourth generation. Specifically, we want to estimate γ_1, γ_2, γ_3 and γ_4 in equation 2.1. The estimation of our model is a challenging task. Ideally, one would like to follow the same individual across different immigrant generations. This is, of course, impossible.

There are several potential sources that contribute to the correlation between G_{it} and u_{it}. First, even when the idiosyncratic term ϵ_{it} is not correlated with G_{it}, there might still be an omitted variable problem due to the unobservable θ_i. The vector θ_i represents the fixed unobserved components that vary at the group level. Different waves of immigration might be characterized by individuals with different potential outcomes. The differential outcomes achieved by individuals belonging to different immigrant generations might then reflect, on top of the true effect of intergenerational assimilation, the differential selection into the various cohorts. Adding controls as proxies for things such as country of origin, family structure and even parental education would probably be not a panacea[5]. Another concern in estimating equation 2.1 using cross-sectional data is that also the assimilation process could

[5]First, the effects of θ_i may be not fully controlled for using such variables. Second, some controls such as parental educational could be a noisy estimate of the true factors, because of the differential quality of institutions across countries.

be cohort specific. That is, immigrants entering the U.S. in different periods could, in principle, exhibit different assimilation trajectories.

A candidate solution to tackle both problems is to follow the same immigrant cohort across different Census years. However, this strategy might still result in biased estimates because of the non-random selection of return migrants. Suppose, for example, that immigrant individuals characterized by adverse traits to succeed in the U.S. economy leave the host country after some time. Using repeated cross-sections, one would then compare the outcome of the average immigrants to the outcome of (the children of) only the best among them.

Figure 2-1 depicts the analysis that we could perform using the NLSY79 or any cross-sectional dataset. It represents the potential outcomes of three different cohorts of immigrants in the U.S.. In the NLSY79, we only observe the outcome of cohort 1 during its first period in the U.S., the outcome of cohort 2 during its second and the outcome of cohort 3 during its third. If we try to infer what will be the second period outcome for cohort 1 by looking at the current outcome of cohort 2, we would add to the true treatment effect for the assimilation into the U.S. society, a noise reflecting the initial starting point, a possibly different speed of assimilation and the selection of return migrants. Only if the covariates that we can use in the NLSY79, such as country of origin, gender, ethnicity, and family structure, completely offset the noise, then we would get unbiased estimates of the true treatment effect. If not, the intergenerational assimilation estimate would be biased, as showed in Figure 2-3. In Figure 2-1, for example, we have assumed that the quality of immigrant cohorts deteriorates over time, and thus the estimates of γ_1 and γ_2 would be too large, as depicted in Figure 2-3.

Combining the NLSY79 Children and Young Adults and the NLSY79 datasets, we can take one step ahead in the estimation of equation 2.1. Specifically, we select the measures of individuals' human capital development that are harmonized in the two datasets and we study the within family variation in measured achievement. We thus define θ_i as a fixed family component. Consider equation 2.1 for an individual

t in the NLSY79 Children and Young Adults that is an immigrant of generation g and his mother, $t-1$, in the NLSY79, immigrant of generation $g-1$. By taking the difference between the equations describing their outcomes, and allowing the constant β_0 to vary between the NLSY79 and the NLSY79 Children and Young Adults datasets, we get

$$y_{it} = Y_{it} - Y_{i(t-1)} = \beta_0^{child} - \beta_0^{mother} + \sum_g \gamma_g \mathbb{1}\{G_{it} = g\} + \beta X_{it} - \beta X_{i(t-1)} + (\epsilon_{it} - \epsilon_{i(t-1)}) \tag{2.2}$$

Using this specification, we can interpret $\beta_0^{child} - \beta_0^{mother}$ as the average difference in achievement between a child and her native mother. γ_g is, then, the difference between the outcome of an immigrant child of generation g relative to his mother' performance, relative to a native child-mother counterpart. This parameter measures the speed of intergenerational assimilation for an immigrant of generation g. Assuming that ϵ_{it} and $\epsilon_{i(t-1)}$ are not correlated with G_{it}, we can estimate equation 2.2 using OLS. We use the following empirical model:

$$y_{it} = \delta_0 + \gamma_2 * C_{it2} + \gamma_3 * C_{it3} + \gamma_4 * C_{it4} + bX_{it} + e_{it} \tag{2.3}$$

where y_{it} is the difference between the outcome of a child and the outcome of his mother, C_{itg} are dummies taking the value of one if the child t is an immigrant of generation g, X_{it} are controls that vary at the individual level, while the error term e_{it} captures child-mother pairs' idiosyncratic shocks.

Combining the NLSY79 Children and Young Adults and the NLSY79, we can therefore make an attempt to isolate the treatment effect of assimilation into the American society. The combination of the Children of NLSY79 and the NLSY79 dataset is crucial, as it allows to follow the same immigrant family over time. We observe the outcomes of two different immigrant generations belonging to the same family. The cost of our empirical strategy, however, is that we can only estimate consistently γ_2, γ_3, and γ_4, but not γ_1. Figure 2-2 depicts the analysis using the combined datasets. Compared to what we could do in Figure 2-1, here we can take into account the bias of estimating the effect of intergenerational assimilation using cross-sectional or repeated cross-sectional data. The resulting estimates are the light blue arrows depicted in Figure 2-3.

The reliability of our empirical strategy holds also when the assimilation of different immigrant cohorts follows different trajectories or when the selection into motherhood is not constant across immigrant generations. In these cases our estimates should be interpreted as average treatment effects on treated individuals. Knowing the extent to which the assimilation process is cohort specific or characterizing the selection into motherhood for each immigrant cohort would be, however, important in evaluating the external validity of our estimates.

2.4 Results

Summary statistics

Table 2.1 provides the summary statistics of the variables we use in the NLSY79. Several patterns emerge in this table. First generation immigrants suffer an educational and cognitive gap compared to natives, but they seem less prone to illicit behaviors. They get 0.5 years less in completed education, while the gaps in mathematics and English are 0.4 and 0.6 of a standard deviation, respectively. Their behaviours are 0.35 of a standard deviation less problematic than those of natives. Furthermore, the results in the cross-section are compatible with immigrant human capital assimilation. Second generation immigrants obtain about 0.4 years of education more and get similar results in cognitive tests, compared to natives. The third generation outperforms other individuals also in cognitive tests, achieving scores higher than those of natives by 0.35 of a standard deviation. In terms of behaviours, second generation immigrants behave as illicitly as natives, while third generation immigrants slightly worse. The presence of both parents in the household is similarly likely across immigrant generations. The patterns are alike when we restrict attention to mothers in the NLSY79, reported in Table B.9. For them, it is interesting to notice that first generation immigrants have about 0.2 to 0.3 more children than native and higher generation immigrant mothers. Furthermore, second and third generation immigrant mothers have the first child 2 years later than native and first generation immigrant mothers.

In the NLSY79 Children and Young Adults we can observe what happens one generation ahead. Table 2.2 shows the summary statistics. There is no gap in educational achievement between children of first generation immigrant mothers and children of native mothers. Also the gap in cognitive skills is smaller than the gap one generation back, being only 0 to 0.1 of a standard deviation for English knowledge and 0.2 for mathematics. Higher order immigrant generations outperform natives in all cognitive tests. In terms of behaviours, children of immigrant mothers behave slightly better than children of native mothers. The gap in behavioural problems between children of first generation immigrant mothers and children of native mothers is about 0.15 of a standard deviation, or half the gap one generation back. In this dataset, the presence of both parents is a bit more likely in immigrant households. As for parenting, children of first generation immigrant mothers have less children than their native counterparts.

Is the dynamics that we observe in the (repeated) cross-section the result of immigrant assimilation across generations? In the next section we use our empirical model to shed some lights on this question.

Main results and heterogeneity

Main results

In Tables 2.3 to 2.7 we show the main results of our analysis. In Table 2.3 we focus on educational achievements. The point estimates for intergenerational assimilation in education between a first generation immigrant mother and her child range from 1.56 to 1.6 extra years, relative to a native counterpart. While the difference in completed years of education between children and their native mothers is about 0.5 years, the gap is about four times larger for a child-mother pair, when the mother is first generation immigrant. It is, instead, not statistically different from 0.5 for child-mother pairs of higher order immigrant generations. The result is robust to the inclusion of several controls, as shown in columns 1 to 3. Furthermore, while the higher educational achievement of children relative to their native mothers is due to a 9 percentage points higher probability of getting a college degree, the result on intergenerational assimilation of children of first generation immigrant mothers is driven by the higher propensity of getting the high school diploma. They are, indeed, about 25 percentage points more likely to complete the high school relative to their mothers, compared to a native child-mother pair. As for higher education, they are 6 percentage points to 9 percentage points more likely than their mothers to get a college degree, relative to a native counterpart, but the estimates are not significantly different from 0. Also for these measures, adding controls to the regression does not affect the estimates on intergenerational assimilation.

In Table 2.4 we show the results on cognitive skills. In both the NLSY79 and the NLSY79 Children And Young Adults datasets we have measures on mathematics and English proficiency. For all these skills, the average native child performs better than his mother. For example, the score in standardized tests obtained by the average child is higher than the score obtained by his native mother by 0.37 of a standard deviation in mathematics and 0.23 to 0.45 in English. The outperformance is even higher for children of first generation immigrant mothers, while, in general, not different from the natives for higher order immigrant generations. In mathematical tests, the child of a first generation immigrant mother obtains a score that is 0.29 of a standard deviation higher than his mother's score, relative to the difference between the scores of a native mother and her child. As for English knowledge, it is between 0.65 and 0.76 of a standard deviation higher than his mother's score, compared to the native counterfactual. The difference between the result on paragraph comprehension and word knowledge is negligible, with the assimilation in terms of word knowledge being larger by 0.1 of a standard deviation. When we add controls for the ethnicity of the child-mother pair, age of mother at birth, age of the child, and number of older siblings, the magnitude of the estimated coefficients reduce by about $\frac{1}{3}$, and the results on mathematics become indistinguishable from 0 in statistical sense.

When we look at the assimilation in terms of attitudes towards antisocial behaviours, in Table 2.5, things are different. First, children in native families have more problematic behaviours than their mothers. The estimated difference is about 0.6 of a standard deviation. Children of first generation immigrant mothers have even worse behaviours compared to their mothers, relative to their native child-mother counterpart. The difference, however, is not significantly different. The children of second and third generation immigrant mothers, instead, behave worse than their mothers, but less so compared to natives. The difference is not significant for children of second generation immigrant mothers, while it is significant and robust to the inclusion of controls for children of third generation immigrant mothers. In particular, the attitude towards antisocial behaviours is 0.23 of a standard deviation less pronounced for children of third generation immigrant mothers, relative to their mothers, compared to a native child-mother pair.

Are the results on schooling, cognitive skills and problematic behaviours driven by differences in family structures? To address this possibility we study how the likelihood of living in two-parent families varies between mothers and children, across different immigrant generations. In Table 2.6 we report the results. The average child in the NLSY79 Children And Young Adults is 5% less likely than his mother to grown up with both parents in the household during his childhood and teenage years. For children of first and third generation immigrant mothers, this difference is substantially smaller in magnitude. However, the difference is not statistically significant. Although there are several potential inputs at the family level that may vary across immigration status and affect the accumulation of skills, our result suggests that the presence of both parents in the household is plausibly not the most important factor. Aside from statistical significance, if the presence of both parents were a key factor for the assimilation in the U.S. society, we would have had similar results in the speed of skill and behavioural assimilation for children of first and third generation immigrant mothers. However, the estimated coefficients described in the previous tables are different across these two groups, for almost all outcome variables and regression specifications.

We then consider some measures of social assimilation. We focus on fertility outcomes and, in particular, on number of children, and, for females, age at first birth and probability of becoming a teenage mother. The results are displayed in Table 2.7. Children of native mothers have on average about 1.6 children less than their mothers. The gap is larger in magnitude, often significantly, for immigrant children. Relative to a native child-mother pair, children of first generation immigrant mothers have 0.54 less children than their mothers. This difference decreases to a non significant 0.17 for children of second generation immigrant mothers, while it is 0.3 for children of third generation immigrant mothers. Adding controls reduce a bit the magnitude of the estimates, but it does not affect the significance of the results for children of first and third generation immigrant mothers. A candidate explanation for this findings is that, as the immigrants assimilate and accumulate human capital specific to the U.S. economy, the relative cost of raising a baby increases. In terms of age at first birth and propensity of becoming a teenage mother, we do not generally find important differences across different immigration status. For example, while the increase in age at first birth for a female child of a first generation immigrant mother, relative to her mother, is twice as large as the increase for a native child-mother counterpart, the coefficient is not significantly different. We only find some significant results on the probability of becoming a teenage mother, when we look at female children of third generation immigrant mothers. While for a female child the probability of giving birth before the age of 18 is 4% less than the probability of her native mother, the decrease reaches 8% to 10% when the mother is a third generation immigrant.

We conclude this section emphasizing that one should be cautious with the interpretation of the result on third generation immigrants and their children. We recognize that our definition of third generation immigrants encompasses different groups of individuals. In particular, mixing is an issue that one should keep in mind in reading our results. For example, in this category there are individuals with only one grandparent born abroad, but also individuals with all grandparents born in a country different from the U.S.. In the NLSY79, unfortunately, we do not observe the country of origin of grandparents other than the paternal grandfather. The effect on assimilation between third generation immigrants and their children - close to zero for most variables - is thus an average effect between the intergenerational assimilation effects of individuals with different number of ancestors born abroad.

Heterogeneity

Our definitions of immigrant generations are quite coarse and each category probably encompasses rather heterogeneous individuals. In the analysis described in the following paragraphs we redefine our immigrant categories along two dimensions. We first differentiate first generation immigrant mothers based on their age at arrival, as well as second generation ones based on the number of parents born abroad. We then split immigrants arrived from Mexico from those arrived from other countries[6].

Let us start with age at arrival and with number of foreign parents. Individuals arrived in the U.S. when they were five years old or younger only attended American schools. Furthermore, among those that started their formal education outside the U.S., those arrived before their teenage years are more likely to cope better with the problems associated to the assimilation into the host country. Similarly, second generation immigrant mothers with only one parent abroad are more likely to be

[6] More than 40% of first and second generation immigrants and about 20% of third generation immigrants in the NLSY79 are from Mexico.

integrated into the U.S. society. As expected, our exercise suggests that the earlier a mother enters the U.S., the more similar the intergenerational dynamics of an immigrant child-mother pair is to the dynamics in a native pair. Furthermore, when a mother is second generation immigrant, with both parents born abroad, the child-mother intergenerational dynamics looks more similar to the dynamics in child-mother pairs with the mother being first generation immigrant. The results of this analysis are shown in Tables 2.8 and 2.9.

We first consider child-mother pairs, with the mother being first generation immigrant. Two main findings emerge. First, child-mother pairs, with the mother arrived in the U.S. when she was five years old or younger, have intergenerational dynamics not distinguishable from the ones of their native counterparts. In a sense, immigrants arrived before the age of six fully exhaust their potential to catch up with the local population. Second, among the other two groups, children of first generation immigrant mothers arrived later are those with the highest results in both educational achievements and cognitive skills tests, relative to their mothers. This is also reflected in the number of children, that is lower relative to the number one generation back, compared to the estimate of a child-mother pair with the immigrant mother arrived earlier. Interestingly, for children of first generation immigrant mothers arrived late, the result on problematic behaviour becomes significant. That is, assimilation into the U.S. society seems to be associated to an increase in the likelihood of showing anti-social attitudes.

Let us now look at child-mother pairs, with the mother being second generation immigrant. At least for some outcomes, including educational achievements, English proficiency, number of children and age at first birth, the intergenerational dynamics of a child-mother pair, with the mother being second generation immigrant born in a family with two immigrant parents, is more similar to the the dynamics of pairs where the mother is first generation immigrant. For most of the coefficients, however, the standard errors are large and the estimates turn out to be insignificant. The only exceptions are the coefficients on high school degree and number of children.

In the first case, children with a second generation immigrant mother and both maternal grandparents born abroad have 7% higher probability of getting the high school diploma relative to their mothers, compared to a native child-mother pair. In the second, they have about 0.4 children less than their mothers, relative to their native counterparts.

Now, instead, we consider immigrants arriving from Mexico separately from other immigrants. The results are shown in Tables 2.10 and 2.11. In terms of schooling, it seems that immigrants from Mexico drive much of the main results described before. For example, children in child-mother pairs with the mother being born in Mexico obtain more than 3 extra years of education, relative to their mothers, compared to their native counterpart. The larger educational achievement is driven by an increased likelihood in obtaining both high school and college degrees, by 44 percentage points and 13 percentage points, respectively. Interestingly, the assimilation expressed in terms of likelihood of getting the high school degree continues also in the following generation. That is, children of second generation immigrant mothers with Mexican origin are 18 percentage points more likely to get the high school diploma, relative to their mothers, compared to their native counterparts. For immigrant child-mother pairs with the mother born in a country different from Mexico, the schooling assimilation is close to zero. Children in child-mother pairs with the mother being first generation immigrant from a country different than Mexico are nonetheless 13 percentage points more likely to get the high school diploma relative to their mothers, compared to their native counterparts. The strong result in English assimilation between first and second generation immigrants, described earlier, is similarly driven by the two immigrant groups. In terms of problematic behaviours, being U.S. born individuals is worse for Mexicans than for immigrants with different origins. Compared to their mothers, children of first generation immigrant mothers from Mexico show behaviours that are 0.44 of a standard deviation worse, relative to a native counterfactual. In this respect, however, it is interesting to notice that female children of first generation immigrant mothers from countries other than Mexico are 8% more likely to give birth as teenagers, compared to their mothers, relative to the native benchmark.

2.5 Further considerations

In the previous section we observed that much of the dynamic assimilation of immigrants exhausts after the second generation. Children of first generation immigrant mothers get levels of cognitive and social skills that look very different compared to their mothers. Higher order generation immigrant families, instead, have an intergenerational dynamics much more similar to the dynamics in native families. This holds for most of the outcome variables we are considering. Are these outcome variables strongly related and thereby the immigrant assimilation the result of a single-dimensional, simple rule? Can we use our analysis to make inference on the evolution of immigrants' selection? In this section we try to address these two questions.

Principal component analysis

To study the interrelation of our different measures of individuals' cognitive and social traits, we use principal component analysis. The results of our analysis are reported in Tables 2.12 and 2.13. As usual with this empirical strategy, we retain the factors that have eigenvalues larger than 1. In Tables 2.12 we use the NLSY79. In Tables 2.13, instead, we use the NLSY79 Children And Young Adults and we express the different measures in terms of deviations from mothers' outcomes.

The result using the NLSY79 indicates that individuals' personality is a multi-faceted and complex object. When we consider schooling outcomes, cognitive skills, behavioural problems and presence of both parents in the household we observe that the two relevant factors account for only about 60% of the whole variance. Furthermore, educational achievements and cognitive skills might be represented by a common factor, but the proxy for illicit behaviours and presence of both parents would be different. When we also consider the parenting outcomes, a third relevant factor emerges. Number of sons, age at first birth and probability of being a teenage mother would be better proxied by this additional factor, rather than those related to education, cognitive skills or problematic behaviours. Our results do not depend on the sample selection. They are similar when we consider the whole NLSY79 sample in Table 2.12 or, as a robustness check, the immigrant sample in Table B.10.

When we use the NLSY79 Children And Young Adults in Tables 2.13, we notice that the child-mother educational gap can be represented by a factor that is not the one representing the cognitive skills gaps. The gaps in behavioural problems or family structure do not share either of the two factors. When we consider female respondents, and we look also at parenting outcomes, an additional factor becomes relevant.

The evolution of immigrants' selection

Our analysis suggests that the development and the success of individuals might be related to several underlying factors. As a result, the catch up along several dimensions that we observe between immigrants and natives is plausibly driven by a multifaceted and complex phenomenon.

Our estimates on the intergenerational assimilation of immigrants in the U.S. suggest that children of first generation immigrant mothers achieve the steady state. In the following generations, no or little improvement occurs, relative to native child-mother pairs. If the intergenerational assimilation dynamics is the same among the different immigrant cohorts in our study, and the propensity to return migration vanishes after the first generation, comparing the outcomes of second and higher order immigrant generations could be a way to get the difference in potential outcomes across the different immigrant cohorts. If so, Tables 2.1 and 2.2 point to a deterioration in the selection of immigrants between the late 1800 / early 1900 and 1980, expressed in terms of potential educational and cognitive skills achievements.

This conclusion, however, does not hold once we keep the area of origin fixed.

Consider Tables 2.14 and 2.15, where we represent the summary statistics of immigrants by area of origin, for the areas with a consistent number of individuals in the NLSY79. In these two Tables, for example, it emerges that second generation immigrants from Mexico are not performing much worse than third generation immigrants of Mexican origin. This is true also for Europeans. For immigrants from Canada the difference across different cohorts is larger, but the sample is quite small. Furthermore, immigrants of Mexican origin tend to underperform relative to those arriving from other areas, irrespective of the immigrant generation. In the Appendix we also look at the country of origin of the different immigrant cohorts in the NLSY79. We find that Europeans account for a large portion of individuals defined as third generation immigrants, but for a small portion of those defined as first or second generation immigrants. In particular, the number of immigrants arriving from Germany, Italy and the UK reduced substantially over time. Mexicans, instead, are a much larger portion of those defined as first and second generation immigrants. Taken together, these results suggest that the deterioration in immigrant potential outcomes is driven by the change in immigrants' origin countries over time.

2.6 Robustness checks

In this subsection we confirm the robustness of our results by performing several tests. First, we repeat the analysis using a different empirical specification. Second, we use different sampling strategies when the outcome variable is educational achievement. Third, we use different measures of reading comprehension and word knowledge.

Different empirical specification: outcomes in levels and family fixed effects

The model in differences that we estimated before would be algebraically equivalent to the model expressed in terms of deviations from means if there were only two individuals per family and one observation per individual. In turn, estimating the model expressed in terms of deviations from means is equivalent to the estimation of a model in levels, with the inclusion of family fixed effects[7]. Our sample consists of mothers with one or more children. Furthermore, for some outcomes we observe multiple observations for the same individual. In this case, therefore, the model in differences and the model in levels with family fixed effects might deliver different estimates. We repeat our analysis and estimate the following model:

$$Y_{it} = \beta_0^{child} + \sum_{g>0} \gamma_g \mathbb{1}\{G_{it} \geq g\} + \beta X_{it} + \theta_i + \epsilon_{it} \tag{2.4}$$

where θ_i is the family fixed effect. In this case, X_{it} only consists of a gender indicator. By avoiding the inclusion of control variables that do not vary across individuals in the same family, such as ethnicity, and that are non-missing for mothers in the NLSY79 dataset, such as age of parents at birth, the results of this model are directly comparable to those in the second columns for each outcome variable in Tables 2.3 to 2.7.

We report the results of this exercise, together with the corresponding ones of Tables 2.3 to 2.7, in Tables B.3 to B.6. With the inclusion of family fixed effects, the magnitude of the coefficients referring to second generation immigrants is a bit smaller, but the significance is unaffected. In this case, the coefficient in the regression of high school graduation corresponding to the indicator for fourth generation immigrants becomes significant. This suggests that reaching the educational steady state of natives might take longer than three immigrant generations. Overall, however, the results are very similar to the ones obtained using the model in differences.

[7]See chapter 5.1 of Angrist and Pischke (2008).

Educational achievements: different sampling strategy

When studying the schooling outcomes in the main section, we adopted some filters to the NLSY79 Children And Young Adults sample, based on the individuals' maximum age at the last interview. We did so to prevent the inclusion of individuals that have not graduated because too young, and not because dropouts. One problem in adopting these filters is that the sample size we obtain is quite small. Another problem is that we might induce some heterogeneity across individuals characterized by different immigration status, along characteristics that correlate with educational achievements.

In Table B.7 we report the results of the analysis using different sampling strategies. The magnitude of the estimated coefficients varies little using different samples, and the significance is unaffected. This exercise suggests that the results on educational achievements described in the main section are not due to the sampling strategy we adopted.

Different measures of reading comprehension and word knowledge

When studying the intergenerational dynamics of cognitive achievements and illicit behaviours, we had to use different measures to capture the development of the same personality trait for mothers and their children. For example, for mothers in the NLSY79 we observe the outcomes in the ASVAB test, while for their children we observe outcomes in the PIAT test. Although both tests measure the development of similar skills, they are different. Similarly, the questions in the NLSY79 leading to the illicit behaviour index are different from those captured in the BPI measure in the NLSY79 Children And Young Adults.

For English ability, however, we observe two measures in both the ASVAB and the PIAT. In the main section, to measure mothers' English ability, we used the combined result, provided by the NLS, in the two English sections of the ASVAB.

Here, we use the fact that the PIAT reading recognition, that measures word recognition and pronunciation ability, and the PIAT reading comprehension, that measures a child's ability to derive meaning from sentences, reflect possibly more closely the skills measured in the ASVAB word knowledge and in the ASVAB paragraph comprehension, respectively. We thus re-estimate our model using the differences between children's and mothers' outcomes in the corresponding sections of the PIAT and ASVAB tests. We report the results in Table B.8. Both the magnitude and the significance of our estimates are unaffected by this exercise.

2.7 Conclusion

In this paper we studied the intergenerational assimilation of immigrants in the U.S.. We measured assimilation along several dimensions of human capital, including educational achievements, cognitive skills, as well as some social attitudes.

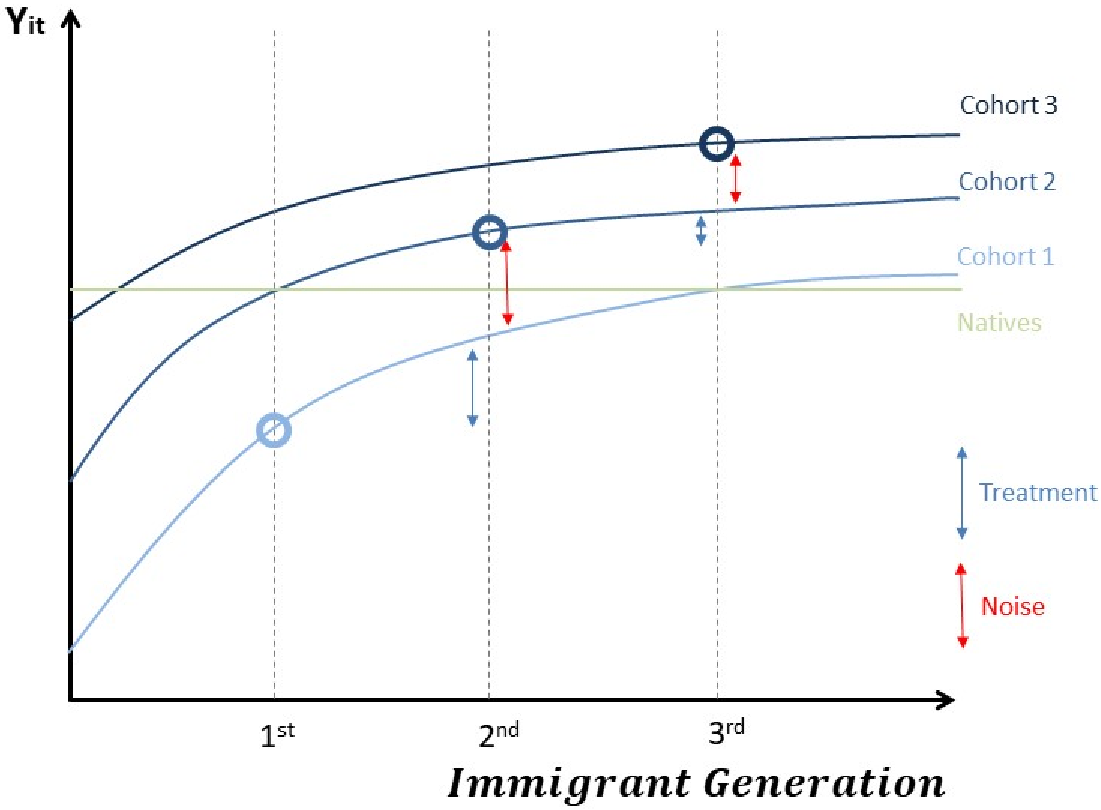
By linking the outcomes of mothers and their children, we made an attempt to separate the treatment effect of intergenerational assimilation from the effect of selection into migration and return migration. We indeed observed two immigrant generations for each immigrant family. The result of our exercise suggests that immigrants take two generations to fully exhaust the potential of educational and cognitive assimilation in the U.S., while it might take longer for other sociability traits. From a policy perspective, our findings suggest that, in order to evaluate the impact of migrants on economies and societies, looking at the outcome of first generation immigrants is not enough. Our results also point to a deterioration in immigrant educational and cognitive potential between the late 1800s / early 1900s and the 1970s, due to a change in origin countries. Individuals from Mexico, indeed, account for a larger portion of the most recent immigrant cohorts, and they systematically suffer in terms of educational and cognitive achievements relative to the other groups.

We also found that not all first and second generation immigrant mothers are the same. We indeed showed that the intergenerational dynamics in child-mother pairs where the mother is first generation immigrant, but with low age at arrival, is similar to the dynamics in native child-mother pairs. Furthermore, the dynamics in child-mother pairs where the mother is second generation immigrant with both parents born abroad is similar to the dynamics in child-mother pairs where the mother is first generation immigrant.

We concluded the paper with a closer look at the foundation of immigrant assimilation. We provided suggestive evidence that the catch up along several dimensions between immigrants and natives is plausibly governed by a multifaceted and complex phenomenon. This implies, for example, that linguistic assimilation might not necessarily lead to immigrant assimilation along all the other personality traits, particularly those related to social skills.

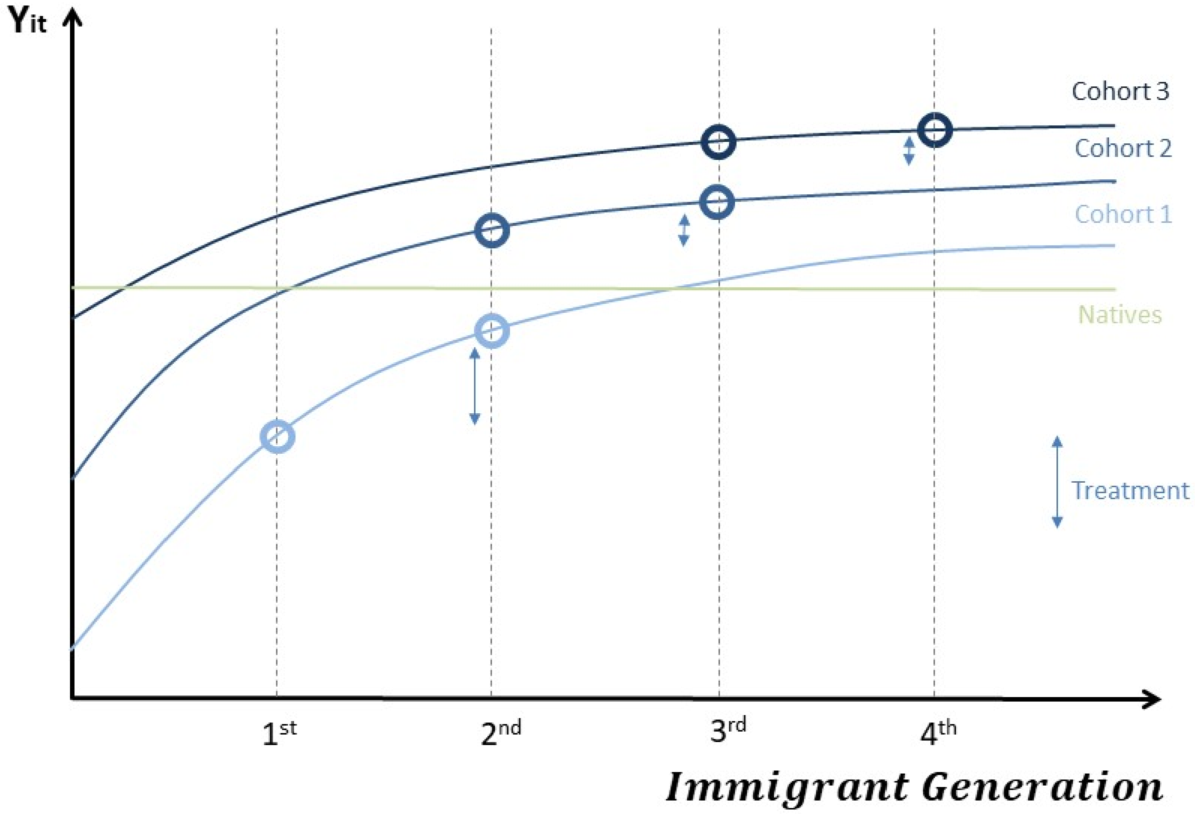
Tables and figures

Figure 2-1: The analysis with cross sectional data



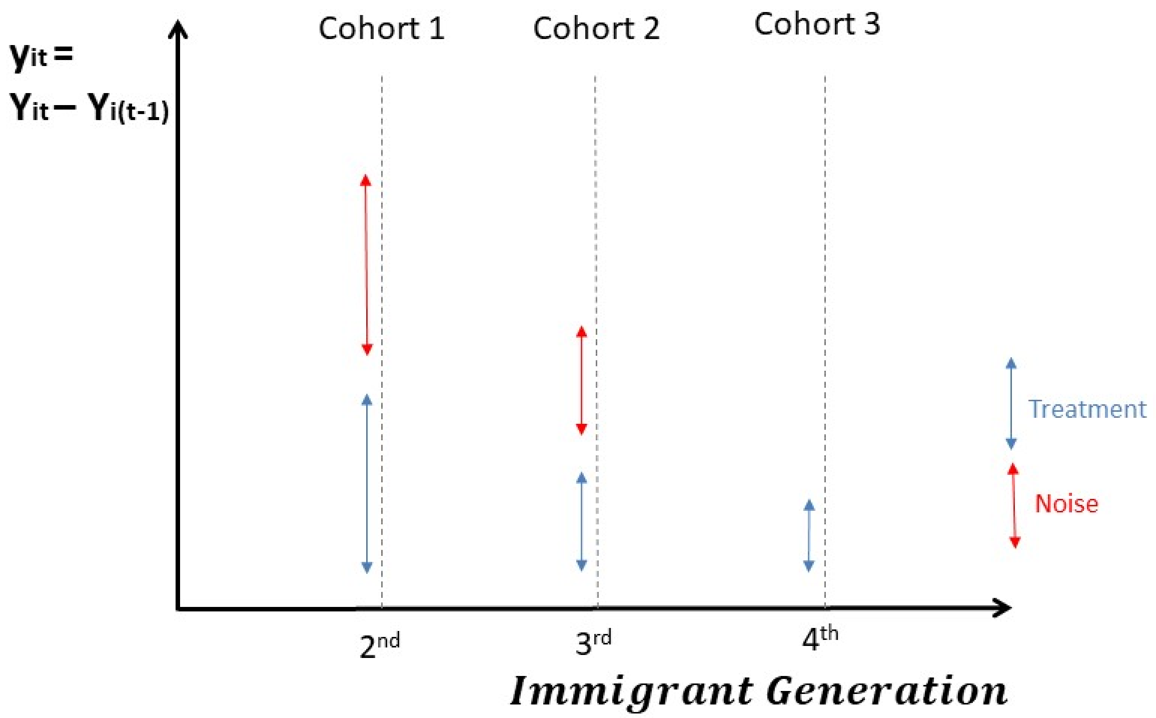
This Figure depicts the analysis that we could perform using the NLSY79 or any cross-sectional dataset. We only observe the outcome of cohorts 1, 2 and 3 at one specific point in time and, specifically, when individuals are first, second and third generation immigrants, respectively. The light blue arrows represent the true dynamics of intergenerational assimilation from one immigrant generation to the next, whereas the red arrows show the noise of the cross-sectional estimates, which is the result of the different starting points of the various cohorts, the different slopes in assimilation, and the selection of return migrants.

Figure 2-2: The analysis with the NLSY79 and the NLSY79 Children and Young Adults



This Figure depicts the analysis that we can perform using the combined NLSY79 and Children of NLSY79. We observe the outcome of cohorts 1, 2 and 3 at two different points in time. Specifically, we observe the outcomes of cohort 1 when individuals belonging to it are first and second generation immigrants; the people in cohort 2 when they are second and third generation immigrants; immigrants in cohort 3 when they are third and forth generation. The light blue arrows represent the true dynamics of intergenerational assimilation from one immigrant generation to the next.

Figure 2-3: Estimates of intergenerational assimilation



This Figure depicts the estimated dynamics of assimilation from one immigrant generation to the next. The light blue arrows represent the true dynamics of intergenerational assimilation from one immigrant generation to the next, whereas the red arrows show the noise of the cross-sectional estimates, which is the result of the different starting points of the various cohorts, the different slopes in assimilation and the phenomenon of return migration.

Table 2.1: Summary statistics; NLSY79

	Native	1st Generation	2nd Generation	3rd Generation
Years of Schooling	13.5	13.06	13.93	14.13
Mathematics Knowledge	-0.02	-0.41	0.06	0.34
English Knowledge	-0.01	-0.66	0.08	0.35
Illicit Behavior	0.00	-0.35	0.00	0.08
Both Parents	0.73	0.77	0.75	0.84
Female	0.50	0.50	0.49	0.46
White	0.80	0.45	0.72	0.92
Black	0.16	0.06	0.02	0.01
Hispanic	0.03	0.49	0.26	0.07
Individuals	10059	537	756	973

Table 2.2: Summary statistics; NLSY79 Children and Young Adults

	Child of Native	Child of 1st Gen.	Child of 2nd Gen.	Child of 3rd Gen.
Years of Schooling	13.44	13.47	13.45	14.00
Mathematics Knowledge	0.24	0.03	0.36	0.47
Reading Comprehension	0.22	0.11	0.31	0.46
Reading Recognition	0.42	0.41	0.54	0.67
Behavioral Problems	0.31	0.15	0.20	0.12
Both Parents	0.58	0.67	0.62	0.70
Sons	1.18	1.06	1.00	0.85
Age at First Birth	21.38	21.89	20.47	22.17
Female	0.48	0.51	0.48	0.52
White	0.77	0.40	0.67	0.88
Black	0.19	0.06	0.04	0.02
Hispanic	0.04	0.53	0.30	0.10
Individuals	8341	522	635	725

Table 2.3: Δ child-mother education

	Δ Years of Schooling			Δ P(High School)			ΔP(College)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Constant	0.53***	0.21**	1.34***	-0.10***	-0.13***	-0.22***	0.09***	0.07***	0.07
	(0.07)	(0.08)	(0.40)	(0.01)	(0.01)	(0.05)	(0.01)	(0.01)	(0.05)
γ_2	1.56***	1.57***	1.60***	0.23***	0.23***	0.24***	0.06	0.06	0.08
	(0.42)	(0.41)	(0.44)	(0.04)	(0.05)	(0.05)	(0.06)	(0.06)	(0.06)
γ_3	-0.17	-0.19	-0.23	0.05	0.05	0.05	-0.02	-0.02	-0.02
	(0.28)	(0.28)	(0.29)	(0.04)	(0.04)	(0.04)	(0.05)	(0.05)	(0.05)
γ_4	-0.07	-0.09	-0.18	0.03	0.03	0.02	-0.01	-0.01	-0.03
	(0.30)	(0.30)	(0.30)	(0.02)	(0.02)	(0.02)	(0.05)	(0.05)	(0.05)
Gender	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Other Controls	No	No	Yes	No	No	Yes	No	No	Yes
N	4378	4378	4378	7103	7103	7103	5308	5308	5308

The Table above reports results of OLS regressions of the difference between a child's and his mother's educational achievement on immigration status. γ_2 refers to the coefficient on the indicator for children whose mothers are first generation immigrants, that is born abroad with parents and paternal grandfather born abroad. γ_3 refers to the coefficient on the indicator for children whose mothers are second generation immigrants, that is born in the U.S. with at least one parent born abroad. γ_4 refers to the coefficient on the indicator for children whose mothers are third generation immigrants, that is born in the U.S. with parents also born in the U.S., but with the paternal grandfather born abroad. *Other Controls* refer to three ethnicity indicators, age of mother at birth and number of older siblings. The sample in the first three columns is restricted to children interviewed at least once at the age of 25 or older. The sample in columns four to six is restricted to children interviewed at least once at the age of 18 or older. The sample in columns seven to nine is restricted to children interviewed at least once at the age of 23 or older. Heteroskedasticity robust standard errors are clustered at the family level and are shown in parenthesis. *, ** , and *** denote significance at the 10%, 5%, 1% level, respectively.

Table 2.4: Δ child-mother cognitive skills

	Δ Maths Knowledge			Δ Reading Comprehension			Δ Word Recognition		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Constant	0.37***	0.38***	0.30***	0.23***	0.18***	1.53***	0.45***	0.37***	0.58***
	(0.02)	(0.02)	(0.08)	(0.02)	(0.02)	(0.08)	(0.02)	(0.02)	(0.08)
γ_2	0.29**	0.29**	0.17	0.65***	0.65***	0.48***	0.76***	0.75***	0.59***
	(0.11)	(0.11)	(0.11)	(0.11)	(0.12)	(0.11)	(0.12)	(0.12)	(0.12)
γ_3	0.10	0.10	0.07	0.04	0.04	0.03	0.08	0.08	0.05
	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.08)	(0.07)	(0.07)	(0.07)
γ_4	-0.00	0.00	0.05	0.02	0.01	0.10*	0.03	0.02	0.10
	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)
Gender	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Other Controls	No	No	Yes	No	No	Yes	No	No	Yes
N	31227	31227	31227	26534	26534	26534	31100	31100	31100

The Table above reports results of OLS regressions of the difference between a child's and his mother's cognitive achievement on immigration status. γ_2 refers to the coefficient on the indicator for children whose mothers are first generation immigrants, that is born abroad with parents and paternal grandfather born abroad. γ_3 refers to the coefficient on the indicator for children whose mothers are second generation immigrants, that is born in the U.S. with at least one parent born abroad. γ_4 refers to the coefficient on the indicator for children whose mothers are third generation immigrants, that is born in the U.S. with parents also born in the U.S., but with the paternal grandfather born abroad. *Other Controls* refer to three ethnicity indicators, age of mother at birth, child's age and number of older siblings. Heteroskedasticity robust standard errors are clustered at the family level and are shown in parenthesis. *, ** , and *** denote significance at the 10%, 5%, 1% level, respectively.

Table 2.5: Δ child-mother problematic behaviours

	Δ Problematic Behaviours		
	(1)	(2)	(3)
Constant	0.58***	0.54***	1.27***
	(0.02)	(0.02)	(0.08)
γ_2	0.15	0.15	0.14
	(0.10)	(0.10)	(0.09)
γ_3	-0.10	-0.10	-0.05
	(0.08)	(0.08)	(0.08)
γ_4	-0.23***	-0.24***	-0.18**
	(0.08)	(0.08)	(0.08)
Gender	No	Yes	Yes
Other Controls	No	No	Yes
N	32787	32787	32787

The Table above reports results of OLS regressions of the difference between a child's and his mother's attitude towards engaging in problematic behaviours on immigration status. γ_2 refers to the coefficient on the indicator for children whose mothers are first generation immigrants, that is born abroad with parents and paternal grandfather born abroad. γ_3 refers to the coefficient on the indicator for children whose mothers are second generation immigrants, that is born in the U.S. with at least one parent born abroad. γ_4 refers to the coefficient on the indicator for children whose mothers are third generation immigrants, that is born in the U.S. with parents also born in the U.S., but with the paternal grandfather born abroad. *Other Controls* refer to three ethnicity indicators, age of mother at birth and number of older siblings. Heteroskedasticity robust standard errors are clustered at the family level and are shown in parenthesis. *, ** , and *** denote significance at the 10%, 5%, 1% level, respectively.

Table 2.6: Δ child-mother presence of both parents in household

	Δ Two-Parents Family		
	(1)	(2)	(3)
Constant	-0.05***	-0.03**	-0.18***
	(0.01)	(0.01)	(0.04)
γ_2	0.05	0.05	0.04
	(0.06)	(0.06)	(0.06)
γ_3	-0.02	-0.02	-0.03
	(0.04)	(0.04)	(0.04)
γ_4	0.04	0.04	0.03
	(0.03)	(0.03)	(0.04)
Gender	No	Yes	Yes
Other Controls	No	No	Yes
N	51076	51076	51076

The Table above reports results of OLS regressions of the difference between a child's and his mother's probability of living in two-parents families on immigration status. γ_2 refers to the coefficient on the indicator for children whose mothers are first generation immigrants, that is born abroad with parents and paternal grandfather born abroad. γ_3 refers to the coefficient on the indicator for children whose mothers are second generation immigrants, that is born in the U.S. with at least one parent born abroad. γ_4 refers to the coefficient on the indicator for children whose mothers are third generation immigrants, that is born in the U.S. with parents also born in the U.S., but with the paternal grandfather born abroad. *Other Controls* refer to three ethnicity indicators, age of mother at birth and number of older siblings. The sample is restricted to children aged 14 or younger. Heteroskedasticity robust standard errors are clustered at the family level and are shown in parenthesis. *, ** , and *** denote significance at the 10%, 5%, 1% level, respectively.

Table 2.7: Δ child-mother parenting

	Δ Number of Kids			Δ Age at First Pregnancy			Δ P(Teenage Pregnancy)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Constant	-1.66***	-1.80***	0.90***	0.95***	0.95***	14.98***	-0.04***	-0.04***	-0.34***
	(0.04)	(0.05)	(0.20)	(0.16)	(0.16)	(0.80)	(0.01)	(0.01)	(0.05)
γ_2	-0.54**	-0.53**	-0.39**	0.94	0.94	1.18	0.01	0.01	0.02
	(0.21)	(0.21)	(0.18)	(0.86)	(0.86)	(0.90)	(0.03)	(0.03)	(0.04)
γ_3	-0.17	-0.17	-0.13	-0.76	-0.76	-0.73	-0.01	-0.01	-0.02
	(0.16)	(0.16)	(0.14)	(0.69)	(0.69)	(0.60)	(0.04)	(0.04)	(0.04)
γ_4	-0.30**	-0.31**	-0.22*	-0.01	-0.01	0.06	-0.04	-0.04	-0.06**
	(0.14)	(0.14)	(0.12)	(0.78)	(0.78)	(0.72)	(0.03)	(0.03)	(0.03)
Gender	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Other Controls	No	No	Yes	No	No	Yes	No	No	Yes
N	5198	5198	5198	1955	1955	1955	3789	3789	3789

The Table above reports results of OLS regressions of the difference between a child's and his mother's parenting behaviour on immigration status. γ_2 refers to the coefficient on the indicator for children whose mothers are first generation immigrants, that is born abroad with parents and paternal grandfather born abroad. γ_3 refers to the coefficient on the indicator for children whose mothers are second generation immigrants, that is born in the U.S. with at least one parent born abroad. γ_4 refers to the coefficient on the indicator for children whose mothers are third generation immigrants, that is born in the U.S. with parents also born in the U.S., but with the paternal grandfather born abroad. *Other Controls* refer to three ethnicity indicators, age of mother at birth and number of older siblings. The sample in the first three columns is restricted to children interviewed at least once at the age of 25 or older. The sample in columns four to six is restricted to female children. The sample the columns seven to nine is restricted to female children interviewed at least once at the age of 18 or older. Heteroskedasticity robust standard errors are clustered at the family level and are shown in parenthesis. *, ** , and *** denote significance at the 10%, 5%, 1% level, respectively.

Table 2.8: Exploring the heterogeneity: different immigrant definitions

	Δ Sch.	Δ P(HS)	ΔP(Col.)	Δ Maths	Δ Read. Comp.	Δ Word Reco.	Δ Probl. Beh.	Δ Two-Par. Fam.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Constant	1.38***	-0.22***	0.07	0.30***	1.53***	0.58***	1.27***	-0.18***
	(0.40)	(0.05)	(0.05)	(0.08)	(0.08)	(0.08)	(0.08)	(0.04)
γ_2, A. E. 0-5	-1.10	0.14	0.05	-0.32	0.08	0.16	0.07	-0.08
	(1.39)	(0.13)	(0.20)	(0.22)	(0.16)	(0.19)	(0.21)	(0.10)
γ_2, A. E. 6-12	0.93***	0.11***	0.15	0.23*	0.28**	0.42***	-0.00	0.06
	(0.35)	(0.04)	(0.11)	(0.12)	(0.12)	(0.14)	(0.14)	(0.07)
γ_2, A. E. 13-17	2.88***	0.38***	0.06	0.41***	0.93***	1.03***	0.29***	0.16
	(0.57)	(0.07)	(0.06)	(0.15)	(0.15)	(0.17)	(0.10)	(0.16)
γ_3, Full	0.48	0.07*	0.01	-0.04	0.15	0.17	-0.07	-0.09*
	(0.40)	(0.04)	(0.07)	(0.11)	(0.12)	(0.13)	(0.12)	(0.05)
γ_3, Half	-0.38	0.04	-0.03	0.12	-0.00	0.02	-0.04	-0.01
	(0.33)	(0.05)	(0.06)	(0.08)	(0.09)	(0.08)	(0.09)	(0.06)
γ_4	-0.17	0.02	-0.03	0.05	0.10*	0.10	-0.18**	0.03
	(0.30)	(0.02)	(0.05)	(0.06)	(0.06)	(0.06)	(0.08)	(0.04)
Gender	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Other Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

N	4378	7103	5308	31227	26534	31100	32787	51076

The Table above reports results of OLS regressions of the difference between a child's and his mother's outcome on immigration status. γ_2 refers to the coefficient on the indicator for children whose mothers are first generation immigrants, that is born abroad with parents and paternal grandfather born abroad. A. E. refers to the mother's age at entry in the U.S.. γ_3 refers to the coefficient on the indicator for children whose mothers are second generation immigrants, that is born in the U.S. with at least one parent born abroad. Full is for mothers with both parents born abroad, Half for mothers with one parent born in the U.S.. γ_4 refers to the coefficient on the indicator for children whose mothers are third generation immigrants, that is born in the U.S. with parents also born in the U.S., but with the paternal grandfather born abroad. *Other Controls* refer to three ethnicity indicators, age of mother at birth and number of older siblings, and, in case of columns four to six, child's age. The sample of each column consists of the same individuals for the corresponding outcomes in the previous tables. Heteroskedasticity robust standard errors are clustered at the family level and are shown in parenthesis. *, ** , and *** denote significance at the 10%, 5%, 1% level, respectively.

Table 2.9: Exploring the heterogeneity: different immigrant definitions

	Δ Number of Kids	Δ Age at First Pregnancy	Δ P(Teenage Pregnancy)
	(1)	(2)	(3)
Constant	0.89***	14.98***	-0.34***
	(0.20)	(0.80)	(0.05)
γ_2, A. E. 0-5	0.07	1.88	0.14
	(0.22)	(3.30)	(0.10)
γ_2, A. E. 6-12	-0.40**	1.29	0.01
	(0.19)	(1.29)	(0.04)
γ_2, A. E. 13-17	-0.58*	0.97	-0.03
	(0.31)	(0.76)	(0.05)
γ_3, Full	-0.41**	0.97	-0.05
	(0.20)	(0.74)	(0.07)
γ_3, Half	-0.06	-1.10*	-0.01
	(0.17)	(0.66)	(0.05)
γ_4	-0.22*	0.06	-0.06**
	(0.12)	(0.72)	(0.03)
Gender	Yes	Yes	Yes
Other Controls	Yes	Yes	Yes

N	5198	1955	3789

The Table above reports results of OLS regressions of the difference between a child's and his mother's outcome on immigration status. γ_2 refers to the coefficient on the indicator for children whose mothers are first generation immigrants, that is born abroad with parents and paternal grandfather born abroad. A. E. refers to the mother's age of entry in the U.S.. γ_3 refers to the coefficient on the indicator for children whose mothers are second generation immigrants, that is born in the U.S. with at least one parent born abroad. Full is for mothers with both parents born abroad, Half for mothers with one parent born in the U.S.. γ_4 refers to the coefficient on the indicator for children whose mothers are third generation immigrants, that is born in the U.S. with parents also born in the U.S., but with the paternal grandfather born abroad. *Other Controls* refer to three ethnicity indicators, age of mother at birth and number of older siblings. The sample of each column consists of the same individuals for the corresponding outcomes in the previous tables. Heteroskedasticity robust standard errors are clustered at the family level and are shown in parenthesis. *, ** , and *** denote significance at the 10%, 5%, 1% level, respectively.

Table 2.10: Exploring the heterogeneity: Mexican vs non-Mexican immigrants

	Δ Sch.	Δ P(S)	ΔP(Col.)	Δ Maths	Δ Read. Comp.	Δ Word Reco.	Δ Probl. Beh.	Δ Two-Par. Fam.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Constant	1.35***	-0.22***	0.08	0.30***	1.53***	0.58***	1.27***	-0.18***
	(0.40)	(0.05)	(0.05)	(0.08)	(0.08)	(0.08)	(0.08)	(0.04)
γ_2^{ME}	3.46***	0.44***	0.13**	0.17	0.49***	0.50***	0.44***	-0.02
	(0.59)	(0.07)	(0.05)	(0.14)	(0.13)	(0.13)	(0.11)	(0.12)
γ_3^{ME}	0.48	0.18***	-0.01	-0.10	0.02	0.01	0.09	-0.15**
	(0.41)	(0.07)	(0.04)	(0.10)	(0.10)	(0.10)	(0.11)	(0.06)
γ_4^{ME}	-0.51	-0.02	-0.02	-0.16	-0.14	-0.19	-0.15	-0.15*
	(0.37)	(0.06)	(0.04)	(0.11)	(0.10)	(0.13)	(0.15)	(0.08)
γ_2^{NO-ME}	0.33	0.13**	0.06	0.12	0.45***	0.60***	0.02	0.04
	(0.47)	(0.06)	(0.09)	(0.14)	(0.14)	(0.16)	(0.11)	(0.07)
γ_3^{NO-ME}	-0.47	0.01	-0.02	0.12	0.02	0.04	-0.07	-0.01
	(0.35)	(0.05)	(0.06)	(0.08)	(0.09)	(0.09)	(0.09)	(0.05)
γ_4^{NO-ME}	-0.10	0.03	-0.03	0.07	0.12**	0.13*	-0.18**	0.04
	(0.33)	(0.02)	(0.06)	(0.07)	(0.06)	(0.07)	(0.08)	(0.04)
Gender	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Other Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

N	4378	7103	5308	31227	26534	31100	32787	51076

The Table above reports results of OLS regressions of the difference between a child's and his mother's outcome on immigration status. γ_2 refers to the coefficient on the indicator for children whose mothers are first generation immigrants, that is born abroad with parents and paternal grandfather born abroad. γ_3 refers to the coefficient on the indicator for children whose mothers are second generation immigrants, that is born in the U.S. with at least one parent born abroad. *ME* refers to Mexican, while *NO-ME* refers to non Mexican immigrants. γ_4 refers to the coefficient on the indicator for children whose mothers are third generation immigrants, that is born in the U.S. with parents also born in the U.S., but with the paternal grandfather born abroad. *Other Controls* refer to three ethnicity indicators, age of mother at birth and number of older siblings, and, in case of columns four to six, child's age. The sample of each column consists of the same individuals for the corresponding outcomes in the previous tables. Heteroskedasticity robust standard errors are clustered at the family level and are shown in parenthesis. *, ** , and *** denote significance at the 10%, 5%, 1% level, respectively.

Table 2.11: Exploring the heterogeneity: Mexican vs non-Mexican immigrants

	Δ Number of Kids	Δ Age at First Pregnancy	Δ P(Teenage Pregnancy)
	(1)	(2)	(3)
Constant	0.90***	14.96***	-0.33***
	(0.20)	(0.80)	(0.05)
γ_2^{ME}	-0.48	0.52	-0.03
	(0.35)	(0.71)	(0.06)
γ_3^{ME}	-0.17	-1.16	0.11
	(0.17)	(0.92)	(0.07)
γ_4^{ME}	0.41**	0.31	0.01
	(0.21)	(1.11)	(0.10)
γ_2^{NO-ME}	-0.26	1.66	0.08*
	(0.17)	(1.40)	(0.04)
γ_3^{NO-ME}	-0.09	-0.47	-0.06
	(0.18)	(0.75)	(0.05)
γ_4^{NO-ME}	-0.30**	0.00	-0.07**
	(0.13)	(0.81)	(0.03)
Gender	Yes	Yes	Yes
Other Controls	Yes	Yes	Yes

N	5198	1955	3789

The Table above reports results of OLS regressions of the difference between a child's and his mother's outcome on immigration status. γ_2 refers to the coefficient on the indicator for children whose mothers are first generation immigrants, that is born abroad with parents and paternal grandfather born abroad. γ_3 refers to the coefficient on the indicator for children whose mothers are second generation immigrants, that is born in the U.S. with at least one parent born abroad. *ME* refers to Mexican, while *NO-ME* refers to non Mexican immigrants. γ_4 refers to the coefficient on the indicator for children whose mothers are third generation immigrants, that is born in the U.S. with parents also born in the U.S., but with the paternal grandfather born abroad. *Other Controls* refer to three ethnicity indicators, age of mother at birth and number of older siblings. The sample of each column consists of the same individuals for the corresponding outcomes in the previous tables. Heteroskedasticity robust standard errors are clustered at the family level and are shown in parenthesis. *, ** , and *** denote significance at the 10%, 5%, 1% level, respectively.

Table 2.12: Principal component analysis; all individuals in NLSY79

	Whole Sample			Mothers			
Cumulative Proportion	0.60			0.59			
	Comp.1	Comp.2	Unexplained	Comp.1	Comp.2	Comp.3	Unexplained
Years of Schooling	0.50		0.21	0.51			0.21
High-School			0.71				0.72
College	0.45		0.37	0.49			0.33
Mathematics	0.49		0.27	0.47			0.31
English	0.48		0.30	0.45			0.35
Illicit Behavior		0.87	0.20			0.80	0.30
Both Parents		-0.41	0.76			-0.54	0.57
Sons	X	X	X		0.60		0.44
Age at First Birth	X	X	X		-0.49		0.37
Teenage Mother	X	X	X		0.55		0.48
Individuals	10234			3825			

The Table above reports results of the principal component analysis using the whole sample of individuals in the NLSY79. We retain factors whose corresponding eigenvalue is larger than 1, and we report factor loadings larger than 0.3.

Table 2.13: Principal component analysis; Δ child-mother in NLSY79 Children And Young Adults

	Whole Sample			Womens				
Cumulative Proportion		0.54				0.64		
	Comp.1	Comp.2	Unexplained	Comp.1	Comp.2	Comp.3	Comp.4	Unexplained
Δ Years of Schooling		0.65	0.17		0.64			0.17
Δ High-School		0.40	0.69		0.33		0.33	0.55
Δ College		0.60	0.33		0.64			0.26
Δ Mathematics	0.51		0.36	0.50				0.37
Δ Reading Comprehension	0.60		0.12	0.60				0.12
Δ Word Knowledge	0.61		0.11	0.60				0.12
Δ Problematic Behavior			0.93				0.76	0.37
Δ Both Parents			0.97				0.54	0.64
Δ Sons	X	X	X			0.44		0.67
Δ Age at First Birth	X	X	X			-0.62		0.31
Δ Teenage Mother	X	X	X			0.62		0.37
Individuals		4405				1476		

The Table above reports results of the principal component analysis using the whole sample of individuals in the NLSY79 Children And Young Adults. The various measures are expressed as the difference between a child's and his mother's outcomes. We retain factors whose corresponding eigenvalue is larger than 1, and we report factor loadings larger than 0.3.

Table 2.14: Summary statistics by area of origin; NLSY79

	Native	1st Generation			2nd Generation			3rd Generation		
Area of Origin		Mexico	Canada	Europe	Mexico	Canada	Europe	Mexico	Canada	Europe
Years of Schooling	13.5	10.9	15.4	13.9	13.0	13.1	14.3	13.1	14.0	14.2
Mathematics Knowledge	-0.02	-0.96	0.38	-0.06	-0.54	-0.02	0.30	-0.41	0.50	0.38
English Knowledge	-0.01	-1.14	0.70	-0.45	-0.52	0.09	0.29	-0.34	0.36	0.39
Illicit Behavior	0.00	-0.52	-0.11	-0.09	-0.09	0.17	-0.00	0.03	-0.07	0.07
Individuals	10059	233	15	82	327	76	226	177	51	613

Table 2.15: Summary statistics by area of origin; NLSY79 Children And Young Adults

	Child of Native	Child of 1st Gen.			Child of 2nd Gen.			Child of 3rd Gen.		
Area of Origin		Mexico	Canada	Europe	Mexico	Canada	Europe	Mexico	Canada	Europe
Years of Schooling	13.4	13.0	15.6	13.6	12.7	12.7	14.5	12.2	10.9	14.3
Mathematics Knowledge	0.24	-0.33	0.40	0.42	-0.17	0.31	0.68	-0.15	0.63	0.54
Reading Comprehension	0.22	-0.28	0.69	0.43	-0.06	0.18	0.53	-0.02	0.59	0.53
Reading Recognition	0.42	-0.04	0.66	0.78	0.14	0.31	0.85	0.11	0.66	0.75
Behavioral Problems	0.31	0.41	-0.32	0.16	0.23	0.41	0.06	0.31	0.36	0.08
Individuals	8341	269	15	60	332	54	132	194	30	429

Chapter 3

A Model of Risk Taking with Experimentation and Career Concerns

3.1 Introduction

Investing in young and innovative firms involves large uncertainty.[1] The Venture Capital financing model offers a solution to deal with the uncertainty inherent to the innovation process: venture capitalists (henceforth VCs) learn about firms over time and hence can condition their financing on the information they acquire. There is large evidence that they differ considerably in their ability to generate returns (see Korteweg and Sorensen (2017)), and that positive past performance by VCs increases their chances to raise a new fund (see the evidence in Kaplan and Schoar (2005)), and the fees they receive from assets under management. Thus, when making their choices they are arguably motivated by career concerns.

Do career concerns prevent VCs from efficiently using their ability to learn about

[1] It has been calculated that around 50% of investments in venture capital exit with zero value, and only about 10% of total investments effectively make all the returns to venture capital vehicles (see Hall and Woodward (2010) and Nanda and Rhodes-Kropf (2013)).

the projects they finance? And which markets are more prone to this problem? In this paper we show that career concerns generally lead to inefficient risk taking. In particular, our novel contribution is to find that the type of experiments that agents can undertake determines the direction of this inefficiency. Moreover, as the number of agents financing projects that share the same idiosyncratic component increases, the inefficiency reduces. In the limit, the equilibrium risk taking approaches the first best.

In the pages that follow we develop a framework where managers can choose between a safe task and a risky one that can be abandoned after an experimentation phase. In our model, both the manager and the market do not know the state of the world, that determines the return on the risky project, and the manager's ability to run the experiment. The manager, independently from his ability, privately receives an initial signal on the state of the world and, based on this, he chooses whether to select the safe or the risky project. If the risky task is chosen, both players observe the binary result of the experimental phase and they abandon the project in case the experiment conveys a bad signal. How informative the experiment is depends on the manager's ability. On average, a good manager extracts better information and produces higher returns from the risky task. However, in some states of the world the good manager might perform worse than a bad one. This happens when the high ability manager receives too often the good signal from the experiment, when it would be better to abandon the risky task, or when he receives too often the bad signal, when the state of the world is positive and the continuation of the risky project would yield high returns.

After characterizing the efficient risk taking rule, we turn the attention to the equilibrium characterization. We first show that every equilibrium features a cutoff strategy: the manager chooses to undertake the risky task if and only if the initial private signal implies that the likelihood of being in the good state of the world is high enough. Also the efficient risk taking rule prescribes a cutoff strategy. However, when we study the welfare properties in our economy, we find that it is intrinsically plagued by inefficient risk taking. The marginal manager, rather than being purely motivated by financial returns, bases his choice also on the wedge between his expected reputation and what the market would think if it knew the true realization of the initial signal. Because sometimes good managers are biased towards abandoning risky projects by mistake, the market could perceive the abandonment of the risky project as a good sign about the managerial ability. If so, in anticipation of the reputational gain that will come from abandonment, also managers that are not particularly optimistic about the state of the world might be induced to choose the risky task. In other circumstances, when abandoning risky projects is perceived by the market as a bad signal about manager's ability, they are inclined to a more prudent behaviour.

We show that our inefficiency result also holds when there are several managers that, upon choosing the same risky task, run independent experiments. We then show that the inefficiency is monotonically decreasing in the number of managers. The intuition of this results builds on the fact that, by observing the outcome of several experiments, the market figures out more often the true state of the world. It becomes, indeed, less and less likely that all managers, correctly in the bad state and by accident in the good one, abandon the risky project. When the market is expected to observe more often the true state of the world, in turn, the wedge between the agent's expected reputation and what the market would think if it knew the realization of the initial signal decreases. This is so because the initial signal is independent on the manager's ability and, thus, it offers no additional information about this ability once the state of the world is observed. Knowing the initial signal, indeed, would be useful to make better inference about the manager's talent only when the risky task is abandoned and the market never gets to know whether it happened correctly or by mistake. We conclude by showing that in the limit the inefficiency disappears.

Relation to the literature. We contribute to three different strands of the literature. First, our paper is related to the recent theoretical literature on the effects of imperfect information about fund managers' abilities. Hochberg et al. (2013) model investors-managers bargaining in a sequential environment where incumbent investors are more informed than outside investors about managers' skills. Marquez et al. (2014), instead, develop a signal-jamming model where fund managers with differential ability to produce returns distorts the fund size decision in order to affect entrepreneurs' learning. Both papers can explain persistence in venture capital funds' returns. Similarly to these models, in our work there is uncertainty about managers skills. However, we focus on how this problem distorts managers' investment decisions once the fund has already been set.

Second, we contribute to the discussion on experimentation in entrepreneurial finance. Recent works, such as and Kerr et al. (2014) and Ewens et al. (2016), emphasize the role of experimentation in nurturing the innovative activity of young firms. We provide a somewhat darker view on the amount of experimentation observed in the venture capital industry. In our model, there can be too much investments in experimental projects.

Third, on a more abstract level, our work is related to the literature on the effect of career concerns on managerial risk taking. In a seminal work, Holmström (1999) shows that when managerial ability directly affects the project success rate and managers care about their reputation, they underinvest in risky projects. A recent paper by Chen (2015) breaks this result by introducing managers' private information on their type and, hence, a signaling motive to take on risk. We maintain, instead, the assumption that managers do not know their ability, but we change the way in which managerial skills affect the returns from undertaking the risky activity. In our modified setting we characterize necessary and sufficient conditions for either type of inefficiency to emerge in equilibrium.

Finally, a setting where agents' learning ability differs in quality improvement - it is the same initially, but not in the intermediate stage - has been modeled by Li

(2007). Unlike in our setting, agents are privately informed about their ability and - unlike in our setting - strategically change their actions as new information arrives. A signalling motive gives them an incentive to give inconsistent reports, similarly to what discarding the project would mean in our setting.

The rest of the paper is organized as follows. Section 3.2 sets up the model with one manager, characterizes the first best and the equilibria under career concerns, and reports their efficiency properties. Section 3.3 extends the analysis to an economy with N managers. Section 3.4 concludes. The proofs that are not in the main text are relegated to the Appendix.

3.2 Baseline model

Setup

Managers, projects, and experiments. An agent, called *manager*, can choose whether to undertake a safe project (S) or a risky project (R). The safe project costs 0, the risky one costs c. The safe one produces returns of v_s - with $v_s > c$ - in any state of the world, while the risky project pays returns v_r - with $v_r > v_s$ - in the good state and nothing in the bad state. Let x denote the state, and the state space be $\mathcal{X} = \{g,\ b\}$. The principal, often referred to as the *market* in this paper, assesses the capability of the manager to anticipate the state of the world.

When the manager chooses the risky task, he runs at no cost an informative experiment to gather additional information about the likelihood of success and, based on the information she gets from the experiment, can decide whether to pursue or to abandon the investment. Let i index the manager's type: a high type ($i = h$) is a manager that is able to extract *better information* from the experiment compared to a low type ($i = l$) in a sense that will become clear in the next lines. We assume that the experiment produces two signals only, denoted s, with $s \in \mathcal{S} = \{g, b\}$. An experiment is then fully described by the precision parameters defining the probability of receiving the *right* signal in each of the two states of the world, $\alpha_i = P(s = g \mid x = g)$ and $\beta_i = P(s = b \mid x = b)$. Notice that superscript i allows signals' precision to differ depending on the manager's type. We further assume that $\alpha_i,\ \beta_i > \frac{1}{2}$ and that parameters are such that it is always optimal to follow the signal. Moreover, the cost of choosing the risky project, c, is only paid when the manager decides not to abandon the project, that is, after she observes a "good" signal ($s = g$), in which case returns realize.[2] Otherwise, in case of abandonment of the risky task because the signal from the experiment is "bad" ($s = b$), the return is zero.

Information and timing. Prior to choosing which project to select, the manager privately observes a signal $\omega \in \mathbb{R}_+$ (which we will refer to as the project's *intrinsic quality*), generated from a density f defined over the support $[\underline{\omega},\ \overline{\omega}]$, which is independent of his type. The manager then updates his prior probability of success of the risky project to $p(\omega)$. Through the analysis we assume that $p'(\omega) > 0$. The signal ω is the only dimension where the manager's and the market's information doesn't coincide.

Players share a common prior belief, ρ, on the probability that the manager is high type. This probability, together with the precision parameters for the two types of agents, ultimately determines the average probabilities of receiving the right signal from the experiment in the two states of the world, that are defined as $\alpha \equiv \rho\alpha_h + (1-\rho)\alpha_l$ and $\beta \equiv \rho\beta_h + (1-\rho)\beta_l$. We denote instead γ the market posterior belief about the manager's type. If the safe project is chosen, returns v_s are realized and the market does not learn anything about the manager's ability. If the risky project is chosen, both the manager and the market observe the realization of the experiment, s. The result of the experiment depends both on the state of the world and on the manager's type; however, conditional on these two pieces of information, it is independent of the realization of the signal ω. The final realization

[2] The net returns are then $v_r - c$ in case of success or $-c$ in case of failure.

of the risky project is also common knowledge and it is the outcome through which the market can update his prior on the manager's ability.[3] We denote γ^l, γ^r the market's posterior when the risky project is pursued and the state of the world is revealed to be good and bad, respectively, and γ^0 the market's posterior if the manager abandons the risky task.

Payoffs. The manager's utility is increasing in profits from the project - which we will call π - and in the market's belief about the probability he is the high type. We will refer to the latter as the career concern motive. We assume the manager is risk-neutral, and that the career concern motive enters linearly in his utility. Specifically, we call the manager's utility $U(\pi, \gamma)$. We assume the following form:

$$U(\pi, \gamma) = (1 - \lambda)\kappa\pi + \lambda\gamma$$

with $\lambda \in [0, 1]$. Under this specification, the parameter λ measures the extent to which the manager is motivated by career concerns, as opposed to maximizing the payoff from the project. κ, the fraction of profits that the manager receives, is an exogenous parameter.

Parameter assumptions. For the sake of simplicity, we assume that parameters are such that it is always optimal to follow the signal, that is, to continue the risky project if and only if the experiment delivers the good signal. Specifically, for this to be true we assume the following:[4]

$$\frac{\alpha_i p(\omega)}{\alpha_i p(\omega) + (1 - \beta_i)(1 - p(\omega))} v_r - c > 0 \quad \forall i, \omega$$

[3] $v_r - c$ or $-c$ when the risky project is not abandoned and the state is good or bad, respectively, or 0 if the signal from the experiment is bad and the risky project is abandoned.

[4] Notice that, as the manager and market do not know the manager's type, this assumption is only a sufficient condition. That is, we might assume as well that for the average manager in the economy would be optimal to follow the result of the experiment, but not for one of the two types of agents.

$$\frac{(1-\alpha_i)\,p\,(\omega)}{(1-\alpha_i)\,p\,(\omega)+\beta_i\,(1-p\,(\omega))}v_r - c < 0 \ \ \forall i, \omega$$

It is evident that the higher α and β are, the more informative the experiment is. However, it might be the case that in a specific environment, detecting a succesfull project is relatively more beneficial than avoiding the loss associated to running a bad one, or viceversa. This will depend on the primitives of the model. We define the high type manager as the one that ensures higher expected profits.

Definition 1 *The manager is high type if for any ω:*

$$p\,(\omega)\,\alpha_h v_r - [(\alpha_h+\beta_h)\,p\,(\omega)-\beta_h]\,c \geq p\,(\omega)\,\alpha_l v_r - [(\alpha_l+\beta_l)\,p\,(\omega)-\beta_l]\,c$$

This condition tells that, for a given ω, the expected return from the risky investment, obtained through the sum of the gain when the manager receives the right signal in the good state, $p\,(\omega)\,\alpha_i(v_r - c)$, and the loss in case the manager gets the wrong signal in the bad state, $-(1-\beta_i)\,(1-p(\omega))\,c$, is larger for the high type manager. Notice that the condition holds when $\alpha_h > \alpha_l$ and $\beta_h > \beta_l$, but could also be satisfied in some cases where the high type receives a more precise signal in one state, but a less precise one in the other state.

Efficient benchmark and equilibrium

In order to make welfare considerations about the equilibrium outcome, we first characterize the efficient project choice, absent any career concerns motive. We then show that, under certain parameter restrictions, every equilibrium is characterized by a cutoff strategy: the manager chooses the risky project if and only if the first period signal, ω, is higher than come cutoff, denoted ω^*.

Through the analysis we call $\sigma : [\underline{\omega}\,;\,\overline{\omega}] \rightarrow [0\,;\,1]$ the manager's mixed strategy;

$\sigma(\omega)$ denotes the probability that the manager chooses the risky project conditional on observing the signal ω.

Efficient benchmark

We define ω_{FB} the signal at which that expected payoffs from the risky and the safe project are equalized. That is, ω_{FB} solves

$$\underbrace{p(\omega_{FB})\alpha(v_r - c) - (1 - p(\omega_{FB}))(1 - \beta)c}_{\pi(risk \mid \omega_{FB})} = \underbrace{v_s}_{\pi(safe \mid \omega_{FB})} \tag{3.1}$$

It is easy to see that, due to the assumption that $p'(\omega) > 0$, the expected returns of the risky project are monotonically increasing in ω. Therefore, efficient project choice prescribes to undertake the risky project if and only if $\omega \geq \omega_{FB}$. Rearranging equation 1, we can characterize the efficient project selection rule as follows.

Remark 2. *The efficient project selection rule is described by:*

$$\sigma(\omega) = \begin{cases} 1 \quad \text{if} \quad \omega \geq \omega_{FB} \equiv p^{-1}\left(\frac{v_s+(1-\beta)c}{\alpha(v_r-c)+(1-\beta)c}\right) \\ 0 \qquad \qquad \text{otherwise} \end{cases} \tag{3.2}$$

Equilibrium

An equilibrium in this economy is a pair specifying the manager's strategy and the principal's posterior about managerial ability, $(\sigma(\omega),\ \gamma)$. Through the text, we consider (Weak) Perfect Bayesian Equilibria. As there might be cases where beliefs are not well defined, we further impose the restriction that players' beliefs are the limiting beliefs computed using totally mixed strategies.[5]

Let us first define and characterize the posteriors on manager's ability that a given strategy profile, $\sigma(\omega)$, would induce. The relevant events, as explained in the previous section, are that a risky project succeeds, fails or is abandoned following the experiment.

Call γ^r, γ^0 and γ^l the posteriors on manager's type when he chooses the risky project, conditional on the project being successful, discarded, or failing, respectively. These are derived in the Appendix and their expressions are given by:

$$\gamma^r \equiv \mathbb{P}(\theta = h \,|\, x = g,\, s = g) = \\ = \rho \, \frac{\int_\omega \mathbb{P}(s = g \,|\, x = g,\, \omega,\, \theta = h)\, p(\omega)\, \sigma(\omega)\, dF(\omega)}{\sum_{i \in \{l,\, h\}} \mathbb{P}(\theta = i) \int_\omega \mathbb{P}(s = g \,|\, x = g,\, \omega,\, \theta = i)\, p(\omega)\, \sigma(\omega)\, dF(\omega)}$$

$$\gamma^0 \equiv \mathbb{P}(\theta = h \,|\, s = b) = \\ = \rho \, \frac{\int_\omega \mathbb{P}(s = b \,|\, \omega,\, \theta = h)\, \sigma(\omega)\, dF(\omega)}{\sum_{i \in \{l,\, h\}} \mathbb{P}(\theta = i) \int_\omega \mathbb{P}(s = b \,|\, \omega,\, \theta = i)\, \sigma(\omega)\, dF(\omega)}$$

$$\gamma^l \equiv \mathbb{P}(\theta = h \,|\, x = b,\, s = g) = \\ = \rho \, \frac{\int_\omega \mathbb{P}(s = g \,|\, x = b,\, \omega,\, \theta = h)\, (1 - p(\omega))\, \sigma(\omega)\, dF(\omega)}{\sum_{i \in \{l,\, h\}} \mathbb{P}(\theta = i) \int_\omega \mathbb{P}(s = g \,|\, x = b,\, \omega,\, \theta = i)\, (1 - p(\omega))\, \sigma(\omega)\, dF(\omega)}$$

The posteriors computed above are clearly affected by the equilibrium strategy profile, since they depend on the set of individuals that choose to undertake the risky

[5] The (Weak) Perfect Bayesian Equilibrium concept would not discipline beliefs in case the risky project is never chosen in equilibrium.

project. The analysis simplifies once we observe that, under some conditions, these sets take a simple interval representation. Notice infact, that since the expected payoff when choosing the risky project is an increasing function of the the realization of the signal ω, as long as career concern motives are not too strong, it is optimal from the manager's perspective to choose the risky project for high values of ω. The following Proposition characterizes the equilibrium strategy.

Proposition 1. *There exists $\tilde{\lambda} \in (0\,;\,1)$ such that, $\forall \lambda < \tilde{\lambda}$, the equilibrium $\sigma(\omega)$ takes the form :*

$$\sigma(\omega) = \begin{cases} 0 & if \quad \omega < \omega^* \\ 1 & if \quad \omega \geq \omega^* \end{cases} \tag{3.3}$$

for some $\omega^* \in [\underline{\omega}; \overline{\omega}]$.

In words, Proposition 1 states that when career concerns are not too strong, *every* equilibrium will exhibit cutoff strategies. In this case, being optimistic enough about the probability of facing a good state of the world ($\omega \geq \omega^*$) is a necessary and sufficient condition for undertaking the risky project ($\sigma(\omega) = 1$). We will refer to the manager receiving the signal ω^* as the *marginal* manager.

Let us now restrict attention to cases where $\lambda < \tilde{\lambda}$. Let now $\tilde{p}$ be defined as the perceived probability of facing the good state of the world, according to the market, once the market observes that a risky project has been chosen. We rewrite the beliefs following the choice of a *risky* project using the cutoff strategy that agents follow in equilibrium. Moreover, in evaluating γ^0, we use the fact that $\mathbb{P}(s = b \,|\, \omega,\, \theta = i) = \mathbb{P}(s = b \,|\, x = g,\, \omega,\, \theta = i)\mathbb{P}(x = g \,|\, \omega,\, \theta = i) + \mathbb{P}(s = b \,|\, x = b,\, \omega,\, \theta = i)\mathbb{P}(x = b \,|\, \omega,\, \theta = i)$ and the definition $\int_{\omega^*}^{\overline{\omega}} \mathbb{P}(x = g \,|\, \omega) \frac{dF(\omega)}{1-F(\omega^*)} \equiv \tilde{p}$. We can then rewrite the posterior beliefs on managerial ability as follows:

$$\gamma^r \equiv \mathbb{P}(\theta = h \,|\, x = g,\, s = g) =$$

$$= \rho \, \frac{\int_{\omega^*}^{\overline{\omega}} \mathbb{P}(s = g \,|\, x = g,\, \omega,\, \theta = h)\, p(\omega)\; dF(\omega)}{\sum_{i \in \{l,\, h\}} \mathbb{P}(\theta = i)\; \int_{\omega^*}^{\overline{\omega}} \mathbb{P}(s = g \,|\, x = g,\, \omega,\, \theta = i)\, p(\omega)\; dF(\omega)} = \rho \frac{\alpha_h}{\alpha}$$

$$\gamma^0 \equiv \mathbb{P}(\theta = h \,|\, s = b) =$$

$$= \rho \, \frac{\int_{\omega^*}^{\overline{\omega}} \mathbb{P}(s = b \,|\, \omega,\, \theta = h)\; dF(\omega)}{\sum_{i \in \{l,\, h\}} \mathbb{P}(\theta = i)\; \int_{\omega^*}^{\overline{\omega}} \mathbb{P}(s = b \,|\, \omega,\, \theta = i)\; dF(\omega)} = \rho \frac{(1-\alpha_h)\tilde{p} + \beta_h(1-\tilde{p})}{(1-\alpha)\tilde{p} + \beta(1-\tilde{p})}$$

$$\gamma^l \equiv \mathbb{P}(\theta = h \,|\, x = b,\, s = g) =$$

$$= \rho \, \frac{\int_{\omega^*}^{\overline{\omega}} \mathbb{P}(s = g \,|\, x = b,\, \omega,\, \theta = h)\, (1 - p(\omega))\; dF(\omega)}{\sum_{i \in \{l,\, h\}} \mathbb{P}(\theta = i)\; \int_{\omega^*}^{\overline{\omega}} \mathbb{P}(s = g \,|\, x = b,\, \omega,\, \theta = i)\, (1 - p(\omega))\; dF(\omega)} = \rho \frac{1-\beta_h}{1-\beta}$$

In general, while every equilibrium features a cutoff strategy, the cutoff is not necessarily unique. This happens because, while the expected payoff from choosing the risky project is increasing with ω, the expected reputation might be a decreasing function of it. In some situations, summarized in the following result, however, we can find sufficient conditions that guarantee a unique equilibrium cutoff.

Corollary 1. *If* $\frac{1-\alpha_h}{\beta_h} > \frac{1-\alpha_l}{\beta_l}$ *and* $\forall \lambda < \tilde{\lambda}$*, the equilibrium cutoff is unique.*

Efficiency

In this section we show that our economy is intrinsically plagued by inefficient managerial investment decisions. We start with a simple example to show where the inefficient risk taking behaviour comes from. We then formalize this example within the more general setting of the model.

A motivating example of preference for risk

Consider a simplified economy where $\omega \in \{0,1\}$, with associated probabilities of good state of the world given by p_0 and p_1. We also assume that parameters are such that if ω is 0 and there were no career concerns, the managers would be indifferent between the safe and the risky project.

In the economy with career concerns, however, the manager with ω equal to 0 strictly prefers the risky project whenever discarding the risky task after the experiment is perceived by the market as a good sign about managerial ability. The manager prefers the risky task whenever

$$(1-\lambda)\kappa\pi(risk \mid \omega = 0) + \lambda E(\gamma \mid \omega = 0)$$

$$>$$

$$(1-\lambda)\kappa\pi(safe \mid \omega = 0) + \lambda\rho$$

where $E(\gamma \mid \omega = 0)$ is the expected reputation from choosing the risky task for the manager observing ω equal to 0. As $\pi(risk \mid \omega = 0) = \pi(safe \mid \omega = 0)$, the only thing that matters is the reputational gain or loss from choosing the risky task. Notice that $E(\gamma \mid \omega_0) = p_0\alpha\rho\frac{\alpha_h}{\alpha} + (1-p_0)(1-\beta)\rho\frac{1-\beta_h}{1-\beta} + (p_0(1-\alpha) + (1-p_0)\beta)\rho\frac{(1-\alpha_h)\tilde{p}+\beta_h(1-\tilde{p})}{(1-\alpha)\tilde{p}+\beta(1-\tilde{p})}$ and that we can rewrite ρ as $p_0\alpha\rho\frac{\alpha_h}{\alpha} + (1-p_0)(1-\beta)\rho\frac{1-\beta_h}{1-\beta} + (p_0(1-\alpha)+(1-p_0)\beta)\rho\frac{(1-\alpha_h)p_0+\beta_h(1-p_0)}{(1-\alpha)p_0+\beta(1-p_0)}$. Simple algebra shows that $E(\gamma \mid \omega = 0) > \rho$, that is the manager is better off choosing the risky task, whenever $\frac{1-\alpha_h}{\beta_h} > \frac{1-\alpha}{\beta}$. When it is so, indeed, it becomes relatively more typical of the best manager to discard, by mistake, the risky task after the experiment in the good state of the world. This creates an incentive to choose the risky task for the manager that is less optimistic than the market ($p_0 < \tilde{p}$). He expects to discard the risky project with higher probability, and he knows that this behaviour will be perceived by the market as the mistake of the high ability manager. This creates a career concern motive for the manager that would be otherwise indifferent between the safe and the risky task.

As long as λ is low enough, every manager in the economy will then choose the risky task. The manager observing ω equal to 0 will do so for reputational reasons, while the manager receiving ω equal to 1 will do so because $\pi(risk \mid \omega = 1) > \pi(safe \mid \omega = 1)$.

Main results

To start with the formalization of our inefficiency results, first recall the expression (3.2), stating that the first best project selection rule requires to undertake the risky project if and only if $\omega \geq \omega_{FB}$. Since in the previous pages we proved that equilibria are characterized by cutoffs, making welfare considerations about the level of risk taking in the economy boils down to comparing the initial signal of the marginal manager in equilibrium, ω^*, to the optimal ω_{FB}. The marginal manager takes on too much or not enough risk depending on the possibility of exploiting a reputational gain or avoiding a reputational loss, by undertaking the risky task. What matters is the wedge between what he and the market will think about his ability in running experiments. If his expected self-assessment is higher than the expectations of the market, the manager will be more cautious and choose the safe project. If, instead, he is less optimistic than the market about the probability of being perceived as high type, he chooses the risky task, even if efficiency requires the safe one. The second possibility arises when discarding risky projects is perceived by the market as a good signal about the managerial ability.

We start with a technical result, that helps to understand how different equilibria induce different market posteriors upon abandoning a risky project. It establishes conditions so that reputation following abandonment is higher, the higher is the market belief on the good state of the world.

Lemma 1. $\gamma^0(\omega^*)$ *is increasing (decreasing) in* ω^* *if and only if* $\frac{1-\alpha_h}{\beta_h} > (<) \frac{1-\alpha}{\beta}$

This result holds because while the threshold ω^* at which the VC is indifferent between the safe and risky project increases, the market becomes more and more optimistic about the state of the world. In this circumstance, a suspension of a risky task is increasingly associated to an error in the good state - which happens, on average, with probability $1-\alpha$ - rather than to a correct forecast when the state of the world is bad - which happens with probability β. The market is thus willing to believe that the manager is a high type when incorrectly discarding is a more salient behaviour of high types rather than low types managers, relative to how often the two correctly abandon the risky project.

As there is a one-to-one correspondence between ω and $p(\omega)$, in the previous Lemma we could have used the probability $p(\omega^*)$ rather than the initial signal on the state of the world, ω^*. Since the marginal manager (the one whose $\omega = \omega^*$) is less optimistic than the market about the state of the world (as $p(\omega^*) < \tilde{p}$) the previous Lemma also suggests that the self-assessment of the marginal manager, upon discarding the risky task, $\rho\frac{(1-\alpha_h)p(\omega^*)+\beta_h(1-p(\omega^*))}{(1-\alpha)p(\omega^*)+\beta(1-p(\omega^*))}$, is lower than the market posterior, $\rho\frac{(1-\alpha_h)\tilde{p}+\beta_h(1-\tilde{p})}{(1-\alpha)\tilde{p}+\beta(1-\tilde{p})}$, if and only if $\frac{1-\alpha_h}{\beta_h} > \frac{1-\alpha}{\beta}$. Viceversa the manager and the market would share the same posteriors in the events of success and failure of the risky project, as these beliefs are independent on the initial signal on the state of the world. The next result, thus, follows:

Lemma 2. *The marginal manager's expectation of his reputation induced by risk taking is higher than the prior if and only if* $\frac{1-\alpha_h}{\beta_h} > \frac{1-\alpha}{\beta}$.

The proof of this Lemma is straightforward. Assume $\frac{1-\alpha_h}{\beta_h} > \frac{1-\alpha}{\beta}$. For the marginal manager,

$$E(\gamma \mid \omega^*) \equiv p(\omega^*)\alpha\rho\frac{\alpha_h}{\alpha} + ((1 - p(\omega^*)(1-\beta))\rho\frac{(1-\beta_h)}{(1-\beta)} +$$

$$+(p(\omega^*)(1-\alpha) + (1-p(\omega^*))\beta)\rho\frac{(1-\alpha_h)\tilde{p} + \beta_h(1-\tilde{p})}{(1-\alpha)\tilde{p} + \beta(1-\tilde{p})}$$

$$>$$

$$p(\omega^*)\alpha\rho\frac{\alpha_h}{\alpha} + ((1 - p(\omega^*)(1-\beta))\rho\frac{(1-\beta_h)}{(1-\beta)} +$$

$$+(p(\omega^*)(1-\alpha) + (1-p(\omega^*))\beta)\rho\frac{(1-\alpha_h)p(\omega^*) + \beta_h(1-p(\omega^*))}{(1-\alpha)p(\omega^*) + \beta(1-p(\omega^*))} = \rho$$

On the left hand side of the inequality we have the expected reputation of the marginal individual, upon observing ω^* and choosing the risky project. On the right hand side we have his expected self-assessment.[6] When the uncertainty about the state of the world is resolved by the manager's action, that is, when the manager gets a good signal in the experiment and pursues the risky project till the end, the manager and the market share the posterior beliefs on the manager's type. In case of success the posterior is $\rho\frac{\alpha_h}{\alpha}$, while it is $\rho\frac{(1-\beta_h)}{(1-\beta)}$ if the project fails. When the manager discards the risky task after observing the outcome of the experiment, however, the players do not know if the project has been interrupted correctly or not. In this scenario, the manager and the market use their different beliefs on the state of the

[6]Notice that since the signal ω is independent of the manager's type, his expected self-assessment - at any ω - equals, of course, the prior, ρ. This is an immediate consequence of the law of iterated expectations.

world, $p(\omega^*)$ and $\tilde{p}$ respectively, to draw a conclusion about the manager's type. The direction of disagreement in this event is disciplined by the condition provided in Lemma 2. When $\frac{1-\alpha_h}{\beta_h} > \frac{1-\alpha}{\beta}$ holds, γ^0 grows with $p(\omega)$, because getting the bad signal by mistake is more typical of an high type. Hence the expected reputation is higher, giving the marginal agent a strict additional gain from taking risk. This means that, in order for him to be indifferent, the project must be worse than the one equalizing monetary payoffs.
With the last Lemma at hand, we are now ready to state our main result providing conditions on the direction of the distortions associated to career concerns.

Proposition 2. *There is over(under) risk-taking in the economy if and only if* $\frac{1-\alpha_h}{\beta_h} > \frac{1-\alpha}{\beta}$ $\left(\frac{1-\alpha_h}{\beta_h} < \frac{1-\alpha}{\beta}\right)$.

To prove this result, we argue that the marginal individual chooses the risky project when the first best would require the safe one. Then we use the fact that any manager receiving the signal ω, where $\omega > \omega^*$, is also taking the risky project. By definition of *first best*, a manager should choose the risky project if and only if $\omega > \omega^{FB}$. At the cutoff ω^{FB}, the expected profit from choosing the *safe* project is identical to the expected profit when choosing the *risky* one, that is

$$(1-\lambda)\kappa v_s = (1-\lambda)\kappa(p(\omega^{FB})\alpha(v_r - c) - (1 - p(\omega^{FB}))(1-\beta)c)$$

By definition, the *marginal* manager is indifferent between the *safe* and the *risky* projects when also the expected reputation is taken into account, that is:

$$(1-\lambda)\kappa v_s + \lambda\rho = (1-\lambda)\kappa(p(\omega^*)\alpha(v_r - c) - (1 - p(\omega^*))(1-\beta)c) + \lambda E(\gamma \mid \omega^*)$$

We put the two conditions together to get:

$$(1-\lambda)\kappa \overbrace{(p(\omega^{FB})\alpha(v_r - c) - (1 - p(\omega^{FB}))(1-\beta)c)}^{\pi(risk \mid \omega^{FB})} + \lambda\rho$$

$$=$$

$$(1-\lambda)\kappa v_s + \lambda\rho$$

$$=$$

$$(1-\lambda)\kappa \underbrace{(p(\omega^*)\alpha(v_r - c) - (1-p(\omega^*))(1-\beta)c)}_{\pi(risk \mid \omega^*)} + \lambda E(\gamma \mid \omega^*)$$

As $E(\gamma \mid \omega^*) > \rho$, it must be the case that $\pi(risk \mid \omega^{FB}) > \pi(risk \mid \omega^*)$. $\pi(risk \mid \omega)$ is an increasing function of ω because $p(\omega)$ is increasing in ω. This is equivalent to $\omega^{FB} > \omega^*$.

Our inefficiency result derives from the wedge between what the manager and the market think about the managerial capability in running experiments, that arises when the risky project is abandoned. Career concerns, in turn, have a bite in the managerial decision problem, because of the expected reputational gains or losses from choosing the risky task. Technically, this gains or losses emerge because the players cannot condition on the state of the world, $x = \{g, b\}$, in evaluating $\gamma^0 \equiv \mathbb{P}(\theta = h \mid s = b)$. In particular, the market conditions this inference on the equilibrium strategy profile, whereas the VC bases it on his observed signal ω. If the counterfactual state of the world in case of a bad draw in the experiment was observed, that is, if players observed what would have happened if the manager continued with the risky task, the expected reputation would coincide with the expected self-assessment, equal to the prior.[7] Similarly, if the VC did not have any

[7] To see this, we just need to compute the posteriors in case players know that the manager was wrong in discarding the risky project, $\rho\frac{1-\alpha_h}{1-\alpha}$, and when the manager was right in doing that, $\rho\frac{\beta_h}{\beta}$. As none of the posterior would now depend on $p(\omega)$ and $\tilde{p}$, the expected reputation and the expected self-assessment would coincide and be equal to $p(\omega)\alpha\rho\frac{\alpha_h}{\alpha} + (1-p(\omega)(1-\beta)\rho\frac{(1-\beta_h)}{(1-\beta)} + p(\omega)(1-\alpha)\rho\frac{1-\alpha_h}{1-\alpha} + (1-p(\omega))\beta\rho\frac{\beta_h}{\beta} = \rho, \ \forall\omega$. This would also be true, in particular, for the

private information on ω, no disagreement on γ^0 would result. In both these cases, no inefficiency would emerge.

Notice that if the *good* manager is better in running experiments in both states of the world, that is $\alpha_h > \alpha_l$ and $\beta_h > \beta_l$, then $\frac{1-\alpha_h}{\beta_h} > \frac{1-\alpha}{\beta}$ never holds. In this scenario, as the following result states, the agents do not take enough risk.

Corollary 2. *If* $\alpha_h > \alpha_l$ *and* $\beta_h > \beta_l$ *there is underinvestment in the risky activity.*

To get, contrary to standard literature, the result of over-investment it must be the case that the high type manager receives a more accurate signal s in bad state of the world, but less accurate one in good state, relative to low type managers. Although Venture Capital is an industry known for some examples of extremely successful investments, it might be characterized by this condition. Some of the most prominent funds, for example, release stories about their failures in detecting successful businesses. This implies that avoiding losses might indeed be as important as detecting successful deals. The most well-known list of failures is the *Anti-Portfolio* by Bessemer Venture Partners, one of the longest-standing venture capital firms. Their list of missed opportunities includes Airbnb, Apple, eBay, Facebook, FedEx, Google and Intel.

Robustness check 1: how robust is it to signalling at the experimentation stage?

In the analysis so far, we have assumed that the outcome of the experiment is public information. This makes it compelling that a manager would follow the informative experiment and continue with the risky project if and only if the signal turns out to be good. However, one might argue that the action at the experimentation stage - whether or not to abandon the project - could be itself a signalling device, in case the experiment outcome is the manager's private information. In this paragraph we argue that, as long as the career concerns motive is not too strong, our main results are robust to adding this additional channel. That is, the unique equilibrium would

marginal manager.

be a separating equilibrium where managers follow the signal.

To see this, let us analyze the subgame where the manager has chosen the risky project, and he has (privately) observed the signal s. If we can show that the equilibrium in this signalling game is one where managers follow the signal, conditional on any ω, then we can conclude that our main results on risk taking behavior still hold. The reason is that in such equilibrium managers would play exactly as they are constrained to do by assumption in our original model.

Let us start focusing on the two possible pooling equilibria. First, consider the case where all managers - independently on s - abandon a project after the experiment. A manager that received $s = g$, would not deviate if and only if:

$$(1-\lambda)*0+\lambda\rho \geq (1-\lambda)\left[\frac{\alpha p(\omega)}{\alpha p(\omega)+(1-\beta)(1-p(\omega))}v_r - c\right]+\lambda\tilde{\gamma}$$

for some induced posterior - $\tilde{\gamma}$ - that is computed by specifying arbitrary off-equilibrium beliefs on s. By the parameters restrictions as in section 2.1, it is easy to observe that the condition can not be satisfied as long as λ is small enough, for any $\tilde{\gamma} \in [0,1]$.

Similarly, consider the case where all managers - independently on s - continue with the project after the experiment. In this case a manager with $s = b$ would not deviate if and only if:

$$(1-\lambda)\left[\frac{(1-\alpha)p(\omega)}{(1-\alpha)p(\omega)+\beta(1-p(\omega))}v_r - c\right]+\lambda\rho \geq (1-\lambda)*0+\lambda\tilde{\gamma}$$

for some $\tilde{\gamma} \in [0,1]$. Again, this is impossible due to the same restrictions, when λ is small enough.

Finally, consider instead a candidate separating equilibrium where the manager continues with the risky project if and only if $s = g$. Here - due to our restrictions - a manager that is solely interested in the material payoffs would want to continue when $s = g$ and abandon when $s = b$. Therefore, since the expected reputation is bounded above by one and below by zero, such equilibrium must exist for some λ small enough.

Robustness check 2: assumption on conditional independence between signals ω and s

In the following lines we show that the assumption on independence between signals ω and s, conditionally on the state of the world, can be optimal from the players' perspective. However, this result does not generalize.

For this exercise, we use a simplified economy where $\omega \in \{0, 1\}$ and we assume that the parameters are such that the manager undertakes the risky task only if ω equals 1. We define $\mathbb{P}(x = g \mid \omega = 1) \equiv p_1$, $\mathbb{P}(s = g \mid x = g, \omega = 1) = \alpha + \epsilon$, and $\mathbb{P}(s = b \mid x = b, \omega = 1) = \beta - \epsilon$, where ϵ captures the degree of positive correlation between signals.

In this case the utility of an agent that, absent career concerns, chooses the risky task upon observing ω equal to 1 is given by

$$U(\omega = 1) = p_1 \Bigg[(\alpha + \epsilon)(v_r - c)\Bigg] + (1 - p_1) \Bigg[(1 - \beta + \epsilon)(-c)\Bigg]$$

As the correlation between signals increases, ϵ increases, and the payoff decreases whenever $c > p_1 v_r$. In this simplified economy, indeed, the independence between signals is better from the players' perspective when the cost of pursuing the risky project after the experiment, c, is high enough relative to the return v_r, or when the probability of success p_1 is low. Put differently, the positive correlation between the signals ω and s becomes detrimental when detecting failing projects during the experimental phase is relatively more important than detecting successful ones.

3.3 An extension: N agents

In this section, we generalize the model by increasing the number of managers to $N > 1$. The structure of the economy is the same of the one analyzed above, although we need some extra assumptions about the timing of the managerial investments, the initial signals ω and the experiments in case of risky project. As for the timing of the economy, we assume that each player observe the final outcome of each investment simultaneously.[8] We also assume that the managers share the same signal ω. With this assumption it follows that either all managers choose the safe project, or they all choose the risky one. As for the experiments, we assume that the realizations are independent across managers, conditionally on the state of the world and their types.

It is easy to observe that, also in the extended model, the efficient rule is the same of the simple model with one manager and also the cutoff strategy in equilibrium, when $\lambda < \tilde{\lambda}$, holds as before. Once again, we focus on this case.

Note that, with N managers, the following three facts become relevant. First, the market and a manager discarding the risky project after the experiment get to know that the state of the world is bad if at least an other manager pursues the investment and this fails. Similarly, they realize that the state of the world is good after suspending the project if there is at least a manager that continues and succeeds. In these two cases their assessment about the managerial ability coincide, as they now condition on the state of the world, rather than on their (different) perceived probabilities of being in either state, pinned down by $p(\omega^*)$ and $\tilde{p}$. Third, even when all managers abandon the risky project after the experiment, and hence there is still uncertainty about the true state of the world, the posterior that market forms is a function of the number of managers in the economy. Indeed, the odds that *all* managers are right or wrong in receiving a bad signal from the experiment differ to the chance that only *one* of them receives the bad signal in either state of

[8] The fact that other managers observe the various outcomes is irrelevant. What matters is that the market observes simultaneously the realizations of the various projects.

the world.

As the manager's type is independent on the signal ω, conditionally on the realization of the experiment and the state of the world, if a manager pursues the risky investment and succeeds the posterior of the market coincides with γ^r; if he fails the posterior equals γ^l. Through this section we rename the posteriors that coincide with the previous analysis as γ_N^r and γ_N^l, respectively. There are other three sets of events inducing different posteriors to the market on a manager's ability. One corresponds to the situation in which the manager abandoned the risky project, but at least another manager pursued it and failed. The second one happens when the manager abandoned the risky project, but at least another manager pursued it and succeeded. Finally, the third refers to the case in which all managers abandoned the risky project. We denote γ_N^{nl}, γ_N^{nr} and γ_N^0 the market posteriors associated to each of these three sets of events. The market's posteriors about the quality of the N^{th} manager, taking as given the performance of the other $N-1$ managers are, therefore, given by:

$$\gamma_N^{nl} \equiv \mathbb{P}(\theta_N = h \,|\, x = b,\, s_N = b) = \rho\,\frac{\beta_h}{\beta}$$

$$\gamma_N^{nr} \equiv \mathbb{P}(\theta_N = h \,|\, x = g,\, s_N = b) = \rho\,\frac{1-\alpha_h}{1-\alpha}$$

$$\gamma_N^0 \equiv \mathbb{P}(\theta_N = h \,|\, s_N = b,\, s_{1,\ldots,N-1} = b) = \rho\,\frac{(1-\alpha_h)(1-\alpha)^{N-1}\tilde{p} + \beta_h\beta^{N-1}(1-\tilde{p})}{(1-\alpha)^N\tilde{p} + \beta^N(1-\tilde{p})}$$

$$\gamma_N^r \equiv \mathbb{P}(\theta_N = h \,|\, x = g,\, s_N = g) = \rho\frac{\alpha_h}{\alpha}$$

$$\gamma_N^l \equiv \mathbb{P}(\theta_N = h \,|\, x = b,\, s_N = g) = \rho\frac{1-\beta_h}{1-\beta}$$

We first establish a result that mirrors the one in Proposition 2.

Proposition 3. *For any finite number of managers, N, the marginal manager takes too much risk if and only if* $\frac{1-\alpha_h}{\beta_h} > \frac{1-\alpha}{\beta}$.

Also in the new economy, there are circumstances in which the market cannot condition the analysis on whether the project was to deliver returns or not, upon observing that the manager got a bad signal from the experiment. Once again, the posterior on manager's ability in this event is the only one where two observers with different opinions on the state of the world would disagree on. As the market is more optimistic on the state of the world compared to the marginal manager, the latter enjoys a reputational benefit or suffers a cost when choosing the risky project, depending on whether the abandonment of it is perceived as a good signal for the market about his quality.

We are interested in comparing economies that differ in N - the number of managers running projects that are linked to the same state of the world. In particular, we want to assess which economies are more plagued by the inefficiency that inevitably results from the pressure of career concerns. To do this, one first observation to be made is that, provided the direction of the inefficiency is the same, it is possible to measure how "strong" is the inefficiency by only looking at how distant is the cutoff associated to the equilibrium marginal manager from the first best value - ω^{FB}. This is stated formally in the next Lemma and further explained in the following lines.

Lemma 3. *Take two economies - denoted 1 and 2 - and associated equilibrium cutoffs ω_1^* and ω_2^*. (i) If, $\frac{1-\alpha_h}{\beta_h} > \frac{1-\alpha}{\beta}$ the expected monetary payoff from economy 1 is higher than in economy 2 whenever $\omega_1^* > \omega_2^*$. (ii) If, $\frac{1-\alpha_h}{\beta_h} < \frac{1-\alpha}{\beta}$ the expected monetary payoff from economy 1 is higher than in economy 2 whenever $\omega_1^* < \omega_2^*$.*

To prove this result, consider the case when $\frac{1-\alpha_h}{\beta_h} > \frac{1-\alpha}{\beta}$. We know in this case equilibria will exhibit excessive risk taking, therefore the cutoffs ω_1^* and ω_2^* would be both lower than ω^{FB}. Assume now $\omega_1^* > \omega_2^*$. We can compare the equilibrium monetary payoffs for each realization of the initial signal ω . To do this, we identify three regions. When $\omega < \omega_2^*$, managers in both economies follow the efficient decision, that is, select the safe project. When $\omega > \omega_1^*$, managers in the two economies take the risky project. When $\omega_2^* \leq \omega \leq \omega_1^*$, managers in economy 2 are selecting the risky project, whereas those in economy 1 choose the safe alternative. Since the first best solution prescribes to select the safe project in these cases, it follows that returns are lower in economy 2 for any realization of ω within this region. Therefore, when taking expectations over all possible values of ω, the monetary payoff is higher in economy 1. The same logic applies to the case when $\frac{1-\alpha_h}{\beta_h} < \frac{1-\alpha}{\beta}$.

The result is useful because it provides us with a simple way to compare different economies in terms of how inefficient is project selection in equilibrium: it is sufficient to establish in which economy the marginal managers departs less from the indifferent one in the first best - absent the career concerns motive. In the following, main result of this section, we show that the inefficiency decreases as the number of managers increases. Economies where N is larger induce equilibria in which expected returns are higher.

Proposition 4. *The inefficiency is monotonically decreasing in the number of managers, N.*

In the proof of this Proposition, we show that $E(\gamma_{N+1} \mid \omega^*) < E(\gamma_N \mid \omega^*)$ if $\frac{1-\alpha_h}{\beta_h} > \frac{1-\alpha_l}{\beta_l}$, while $E(\gamma_{N+1} \mid \omega^*) > E(\gamma_N \mid \omega^*)$ whenever $\frac{1-\alpha_h}{\beta_h} < \frac{1-\alpha_l}{\beta_l}$. That is, the marginal manager in the economy with $N+1$ managers finds less appealing to invest in the risky activity from a career perspective, compared to the same individual in the economy with N managers, exactly when there is a reputational gain by choosing it. Instead, the risky task is now becoming more appealing when it is associated to a reputational disadvantage. This clearly implies that the bite of career concerns is more loose and that the marginal individual is characterized by a signal ω^* closer to the first best cutoff, ω^{FB}.

The following Proposition states that in the limit the inefficiency disappears.

Proposition 5. *As the number of managers, N, goes to infinity, the inefficiency disappears. That is, ω^* approaches ω^{FB}.*

The proof of this result is very simple. Consider the expected reputation of the marginal manager, that we now denote $E(\gamma_N \mid \omega^*)$:

$$E(\gamma_N \mid \omega^*) \equiv p(\omega^*)\alpha\rho\frac{\alpha_h}{\alpha} + ((1 - p(\omega^*)(1-\beta))\rho\frac{(1-\beta_h)}{(1-\beta)} +$$

$$+p(\omega^*)(1-\alpha-(1-\alpha)^N)\rho\frac{1-\alpha_h}{1-\alpha} + (1-p(\omega^*))(1-(1-\beta)-\beta^N)\rho\frac{\beta_h}{\beta} +$$

$$+(p(\omega^*)(1-\alpha)^N + (1-p(\omega^*))\beta^N)\rho\,\frac{(1-\alpha_h)(1-\alpha)^{N-1}\tilde{p} + \beta_h\beta^{N-1}(1-\tilde{p})}{(1-\alpha)^N\tilde{p} + \beta^N(1-\tilde{p})}$$

As $1-\alpha$ and β are numbers in the interval [0, 1], $(1-\alpha-(1-\alpha)^N)$ and $(1-(1-\beta)-\beta^N)$ tend to $(1-\alpha)$ and β, respectively. Furthermore, since the posterior belief associated to the event in which all managers suspend the investment in the risky project, $\frac{(1-\alpha_h)(1-\alpha)^{N-1}\tilde{p}+\beta_h\beta^{N-1}(1-\tilde{p})}{(1-\alpha)^N\tilde{p}+\beta^N(1-\tilde{p})}$, is also in the interval [0, 1], and the weight on this posterior approaches 0, we have that the following result:

$$\begin{aligned}\lim_{N\to\infty} E(\gamma_N \mid \omega^*) \;&=\; p(\omega^*)\alpha\rho\frac{\alpha_h}{\alpha} + ((1-p(\omega^*)(1-\beta))\rho\frac{(1-\beta_h)}{(1-\beta)} + \\ &\quad + p(\omega^*)(1-\alpha)\rho\frac{1-\alpha_h}{1-\alpha} + (1-p(\omega^*))\beta\rho\frac{\beta_h}{\beta} \;=\; \rho\end{aligned}$$

This implies that the marginal manager has no reputational gain in expectations by choosing the risky task. Therefore ω^* and ω^{FB} must coincide and the inefficiency disappears.

The intuition behind this Proposition is straightforward. When the number of managers tends to infinity, the chance for the market not to observe the counterfactual state of the world, that occurs when all managers abandon the risky project, approaches zero. This is so because, when the state of the world is good, it is very unlikely that every manager has, incorrectly, a bad draw in the experiment. In contrast, when the state of the world is bad, it is very likely that at least one agent receives by mistake the signal to pursue the investment. In the limit, therefore, the wedge between the managers' self assessments and their expected reputations vanishes.

Robustness check: N agents with conditionally independent signals ω

In this section, we generalize the model by increasing the number of managers to $N > 1$, relaxing the assumption of commonality of signals ω for all managers. We now assume that signals ω are independent across managers, conditionally on the state of the world. As for the experiments, we retain the assumption of the simplest extension, that is we assume that the realizations are independent across managers, conditionally on the state of the world and their types.

The details on the results of this extension are in the Appendix. Once again, the efficient rule requires the selection of the risky project if and only if the signal ω is high enough. With conditionally independent ω, however, it is possible that some managers choose the risky project, while others select the safe one. Relative to the original extension, there are now several events in which the market cannot observe the counterfactual state of the world. As before, this happens when no manager pursues the risky project after observing the outcome of the experiment. However, in this economy the number of experiments does not necessarily equal the number of managers, since some managers might choose the safe project, while the others opt for the risky one. In assessing the counterfactual state of the world, the market makes use of the information on how many managers received a signal ω inducing the choice of the risky project.

In this economy we need to define N+4 posteriors. One relates to the scenario in which the N^{th} manager pursues the risky investment and succeeds. In this case the posterior of the market coincides with γ_N^r. One in which the manager pursues the risky investment and fails. In this case the posterior of the market coincides with γ_N^l. One corresponds to the situation in which the manager abandons the risky project, but at least another manager pursues it and fails. In this scenario the posterior is given by γ_N^{nl}. One is relevant when the manager abandons the risky project, but at least another manager pursues it and succeeds. In this scenario the posterior is given by γ_N^{nr}. Then, there are N cases in which the manager abandons

the risky project following the experiment and the market cannot be certain about the state of the world. This happens when the other N-1 managers choose the safe project, or when the other N-1 abandons the risky project after the experiment, or in the other N-2 possible cases in which some but not all managers choose the safe project while all the others discard the risky one after the experiment.
As we show in details in the Appendix, the economy is characterized by the same sort of inefficiency described in the simplest extension. Furthermore, as the number of managers tends to infinity the inefficiency disappears. As in the simplest extension, when the number of managers tends to infinity, the chance for the market not to understand the underlying state of the world approaches zero, when evaluating a manager discarding the risky task following the experiment. Here, this happens despite the possibility that most of the managers undertake the safe project. In the limit, indeed, the set of managers choosing the risky task becomes large. This makes very unlikely that, incorrectly when the state of the world is good, each manager in this set receives a bad signal from the experiment. In contrast, when the state of the world is bad, it becomes very likely that at least one manager in this set receives by mistake the signal to pursue the investment. As in our simplest extension, in the limit, the wedge between the manager's self assessment and his expected reputation vanishes. Thus each manager chooses the project according to the first best rule.

3.4 Conclusion

In this Chapter we proposed a setting where information about a manager's ability is imperfect and managers are interested in their reputation. Motivated by the application to investments in young firms, we modeled managers as agents that create value because they can experiment and learn about a projects potential. As it is greatly emphasized by, among others, Kerr et al. (2014), the ability to learn about a project's profitability at relatively early stages is a skill that venture capitalists must have in order to succeed in the industry. Infact, experimentation is desirable to the extent that it provides the incentive to finance innovative and young firms. However, it is reasonable to think that some venture capitalists are better at extracting information from early experiments than others. If this skill is so important, then naturally venture capitalists would benefit from making investors and entrepreneurs believe that they are good in this dimension. It would increase their bargaining power, and help them find better deals at the fundraising stage. In light of this observation, we studied venture capitalists' incentive to take on risk when career concerns are at play, that is, outside observers are learning about their ability to experiment. Contrary to Holmström (1999), where managers add value because they directly increase a project's success rate and in equilibrium they become too risk-averse, agents in this model might take inefficiently high risk. The reason is that the abandonment of a promising project at an intermediate stage might be good news about the agent's ability. In particular, this is the case when a good venture capitalist is typically one whose experimenting technology is biased towards receiving negative outcomes. In this situation a venture capitalist would tend to opt excessively for risky, experimental business strategies, in anticipation of the reputational gain that comes from cutting off the investment at an intermediate stage. This result provides a somewhat darker point of view on the amount of experimentation and risk observed in the industry. We studied one solution to this problem in one extension of the model, where we show that when the observer gets information about the outcome of similar projects, the inefficiency is reduced. The reason is that what drives the inefficiency is the fact that the observer can't distinguish whether a project was abandoned because doomed to fail or due to a false negative in the experiment. The information from similar projects provides the observer with an imperfect signal about the counterfactual. It is usually argued that when agents interact less frequently, career concerns are more severe. In this context, one would expect younger VC firms' decision to be particularly distorted. The novel empirical implication of our model is that the markets more plagued by career concerns distortions are those where agents engage in more unique and less correlated activities.

Appendix A

Appendix to chapter 1: Do not Put Off Until Tomorrow What You Can Do Today: Age at Arrival and Immigrants' Human Capital

The model

I assume that the world is a two-period economy: in the first period parents choose the timing of migration to the U.S., in the second one immigrant children select their investment in human capital. Parents are indexed with i, whereas immigrant children with t.

I start from the parental problem. To simplify the analysis, I assume that the only variable that parents can choose is the timing of migration to the U.S., t_i, while all the other parameters are treated as exogenous. Parents' utility is governed by the function U_i, which is affected by the educational achievement of their N_i children, $e_t(t_i)$, as well as by the timing of migration to the U.S.. In solving for the optimal timing of migration, parents also take into account the extra cost of raising their children in the U.S.. The parental problem is described by

$$\max_{t_i} \quad U_i\big(y_1(t_i), ..., y_{N_i}(t_i), t_i\big) - \sum_{t=1}^{N_i} \big(A_i - (t_i - b_{it})\big) \tag{A.1}$$

where A_i is the children' age at which they stop being dependent from their parents, whereas b_{it} is the year of birth of child t in family i.

It is then easy to derive the optimal condition. The optimal timing of migration is implicitly defined by

$$\sum_{t=1}^{N_i} \frac{\delta U_i}{\delta y_t} \frac{\delta y_t}{\delta t_i}(t_i^*) + \frac{\delta U_i}{\delta t_i}(t_i^*) + N_i = 0 \tag{A.2}$$

Conditionally on the exogenous year of birth of individual t, there is a function between parents i and the age at arrival in the U.S. of their children, $a^*_{it} \equiv t^*_i - b_{it}$. In the rest of the analysis I will call this function $\Phi : i \mapsto a$. To simplify the characterization for the relation between investment in human capital and age at arrival, I will assume that the function Φ is one-to-one.

In the second period, immigrant children observe their age at arrival in the U.S. and they choose their investment in human capital. I define $v(y)$ the utility they get from a stock y of human capital, while I assume that its cost has two components: the first depends on the age at arrival, $c(y, a)$, and the second is linear in a family specific component, $\mu(i)$. The problem of child t in family i is described by

$$\max_{y} \quad v(y, a) - c(y, a) - \mu(i)y \tag{A.3}$$

Using the function linking families to age at arrival in the U.S., $\Phi(i)$, I can rewrite the problem as

$$\max_{y} \quad v(y, a) - c(y, a) - \mu\big(\Phi^{-1}(a)\big)y \tag{A.4}$$

The optimal investment in human capital satisfies the following equation:

$$v_y(y^*, a) = c_y(y^*, a) + \mu\big(\Phi^{-1}(a)\big) \tag{A.5}$$

where v_y and c_y are the marginal utility and the part of marginal cost, that, conditionally on a, is independent on the family component. Using total differentiation I can then easily get the change in human capital associated to a unit change in age of arrival in the U.S.. This is given by

$$\frac{dy}{da} = \frac{c_{ya} - v_{ea}}{v_{yy} - c_{yy}} + \frac{\frac{\delta\mu}{\delta i}\frac{1}{\frac{\delta\Phi}{\delta i}}\left(\Phi^{-1}(a)\right)}{v_{yy} - c_{yy}} \tag{A.6}$$

where c_{ya}, c_{yy} and v_{yy} are, respectively, the cross derivative of the part of the educational cost which is independent on family characteristics, its second derivative with respect to human capital and the second derivative of the utility with respect to human capital. The first part of the equation represents what I refer to as the causal effect of age at arrival on skill accumulation, while the second one is the selection term.

Robustness checks

Presence of both parents in households

Table A.1: The effect of age at arrival on years of education; family fixed effects regressions controlling for presence of both parents

Sample	Full				Same Year of Arrival			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Age At Arrival	-0.14**		-0.14		-0.16*		-0.20	
	(0.06)		(0.09)		(0.08)		(0.13)	
Age At Arrival 6-8		0.11		0.12		-0.32		-0.39
		(0.45)		(0.50)		(0.47)		(0.53)
Age At Arrival 9-11		-0.92*		-0.91*		-1.40**		-1.50**
		(0.49)		(0.54)		(0.64)		(0.72)
Age At Arrival 12-14		-1.32**		-1.29		-1.74**		-1.95**
		(0.58)		(0.78)		(0.71)		(0.97)
Age At Arrival 15-17		-1.66**		-1.62		-2.05**		-2.32*
		(0.72)		(0.98)		(0.83)		(1.18)
Family F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Both Parents	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Birth Order	No	No	Yes	Yes	No	No	Yes	Yes
N	256	256	256	256	213	213	213	213

The Table above reports the results of regressions of years of education on immigrants' age at arrival in the U.S.. In odd columns I report the results of a linear regression model, whereas even columns display the results of a model using categorical variables for intervals of age at arrival in the U.S.. In each specification I control for a sex indicator, a dummy indicator for the presence of both parents in the household, family fixed effects and, when specified, a birth order control. In the first four columns the sample consists of immigrant individuals arrived in the U.S. at the age of 17 or younger, with at least one sibling in the sample. In columns five to eight the sample consists of immigrant individuals arrived in the U.S. at the age of 17 or younger, with at least one sibling in the sample, with the additional requirement that each sibling arrived in the U.S. in the same calendar year. Heteroskedasticity robust standard errors are clustered at the family level and they are shown in parenthesis. *, ** , and *** denote significance at the 10%, 5%, 1% level, respectively.

Table A.2: The effect of age at arrival on word knowledge; family fixed effects regressions controlling for presence of both parents

Sample	Full				Same Year of Arrival			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Age At Arrival	-0.06***		-0.05*		-0.05**		-0.04	
	(0.02)		(0.03)		(0.02)		(0.03)	
Age At Arrival 6-8		-0.02		-0.01		-0.06		-0.05
		(0.14)		(0.15)		(0.13)		(0.15)
Age At Arrival 9-11		-0.18		-0.15		-0.15		-0.13
		(0.11)		(0.14)		(0.10)		(0.16)
Age At Arrival 12-14		-0.51***		-0.48*		-0.49***		-0.50**
		(0.19)		(0.25)		(0.18)		(0.24)
Age At Arrival 15-17		-0.63**		-0.57*		-0.54**		-0.54
		(0.25)		(0.33)		(0.25)		(0.35)
Family F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Both Parents	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Birth Order	No	No	Yes	Yes	No	No	Yes	Yes
N	4511	4511	4511	4511	4476	4476	4476	4476

The Table above reports the results of regressions of word knowledge on immigrants' age at arrival in the U.S.. In odd columns I report the results of a linear regression model, whereas even columns display the results of a model using categorical variables for intervals of age at arrival in the U.S.. In each specification I control for a sex indicator, a dummy indicator for the presence of both parents in the household, family fixed effects and, when specified, a birth order control. All these controls are interacted with a dummy variable indicating the immigration status. The sample in each column consists of native and immigrant siblings. In the immigrant sample of the first four columns I keep individuals arrived in the U.S. at the age of 17 or younger, with at least one sibling in the sample. In columns five to eight the immigrant sample consists of immigrant individuals arrived in the U.S. at the age of 17 or younger, with at least one sibling in the sample, with the additional requirement that each sibling arrived in the U.S. in the same calendar year. Heteroskedasticity robust standard errors are clustered at the family level and they are shown in parenthesis. *, ** , and *** denote significance at the 10%, 5%, 1% level, respectively.

Table A.3: The effect of age at arrival on paragraph comprehension; family fixed effects regressions controlling for presence of both parents

Sample	Full				Same Year of Arrival			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Age At Arrival	-0.07**		-0.09**		-0.07**		-0.12**	
	(0.03)		(0.04)		(0.03)		(0.06)	
Age At Arrival 6-8		0.15		0.05		0.08		-0.05
		(0.16)		(0.17)		(0.16)		(0.18)
Age At Arrival 9-11		-0.18		-0.30		-0.18		-0.39
		(0.17)		(0.19)		(0.21)		(0.25)
Age At Arrival 12-14		-0.69***		-1.02***		-0.70***		-1.17***
		(0.23)		(0.28)		(0.26)		(0.34)
Age At Arrival 15-17		-0.85***		-1.27***		-0.90***		-1.50***
		(0.29)		(0.37)		(0.32)		(0.42)
Family F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Both Parents	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Birth Order	No	No	Yes	Yes	No	No	Yes	Yes
N	4511	4511	4511	4511	4476	4476	4476	4476

The Table above reports the results of regressions of paragraph comprehension on immigrants' age at arrival in the U.S.. In odd columns I report the results of a linear regression model, whereas even columns display the results of a model using categorical variables for intervals of age at arrival in the U.S.. In each specification I control for a sex indicator, a dummy indicator for the presence of both parents in the household, family fixed effects and, when specified, a birth order control. All these controls are interacted with a dummy variable indicating the immigration status. The sample in each column consists of native and immigrant siblings. In the immigrant sample of the first four columns I keep individuals arrived in the U.S. at the age of 17 or younger, with at least one sibling in the sample. In columns five to eight the immigrant sample consists of immigrant individuals arrived in the U.S. at the age of 17 or younger, with at least one sibling in the sample, with the additional requirement that each sibling arrived in the U.S. in the same calendar year. Heteroskedasticity robust standard errors are clustered at the family level and they are shown in parenthesis. *, ** , and *** denote significance at the 10%, 5%, 1% level, respectively.

Table A.4: The effect of age at arrival on arithmetic reasoning; family fixed effects regressions controlling for presence of both parents

Sample	Full				Same Year of Arrival			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Age At Arrival	-0.02		-0.05		-0.01		-0.06	
	(0.03)		(0.03)		(0.04)		(0.05)	
Age At Arrival 6-8		-0.10		-0.19		-0.04		-0.16
		(0.12)		(0.15)		(0.13)		(0.18)
Age At Arrival 9-11		-0.12		-0.22		-0.04		-0.21
		(0.16)		(0.22)		(0.20)		(0.29)
Age At Arrival 12-14		-0.21		-0.53^{**}		-0.14		-0.58^{*}
		(0.26)		(0.26)		(0.27)		(0.33)
Age At Arrival 15-17		-0.20		-0.60		-0.16		-0.71
		(0.29)		(0.40)		(0.31)		(0.49)
Family F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Both Parents	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Birth Order	No	No	Yes	Yes	No	No	Yes	Yes
N	4511	4511	4511	4511	4476	4476	4476	4476

The Table above reports the results of regressions of arithmetic reasoning on immigrants' age at arrival in the U.S.. In odd columns I report the results of a linear regression model, whereas even columns display the results of a model using categorical variables for intervals of age at arrival in the U.S.. In each specification I control for a sex indicator, a dummy indicator for the presence of both parents in the household, family fixed effects and, when specified, a birth order control. All these controls are interacted with a dummy variable indicating the immigration status. The sample in each column consists of native and immigrant siblings. In the immigrant sample of the first four columns I keep individuals arrived in the U.S. at the age of 17 or younger, with at least one sibling in the sample. In columns five to eight the immigrant sample consists of immigrant individuals arrived in the U.S. at the age of 17 or younger, with at least one sibling in the sample, with the additional requirement that each sibling arrived in the U.S. in the same calendar year. Heteroskedasticity robust standard errors are clustered at the family level and they are shown in parenthesis. *, ** , and *** denote significance at the 10%, 5%, 1% level, respectively.

Table A.5: The effect of age at arrival on numerical operations; family fixed effects regressions controlling for presence of both parents

Sample	Full				Same Year of Arrival			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Age At Arrival	-0.05		-0.04		-0.05		-0.06	
	(0.04)		(0.05)		(0.04)		(0.06)	
Age At Arrival 6-8		-0.02		-0.03		-0.15		-0.19
		(0.27)		(0.29)		(0.29)		(0.31)
Age At Arrival 9-11		-0.07		-0.08		-0.22		-0.27
		(0.31)		(0.33)		(0.37)		(0.40)
Age At Arrival 12-14		-0.52		-0.55		-0.72*		-0.85
		(0.36)		(0.42)		(0.42)		(0.52)
Age At Arrival 15-17		-0.41		-0.45		-0.49		-0.65
		(0.54)		(0.62)		(0.58)		(0.71)
Family F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Both Parents	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Birth Order	No	No	Yes	Yes	No	No	Yes	Yes
N	4511	4511	4511	4511	4476	4476	4476	4476

The Table above reports the results of regressions of numerical operations ability on immigrants' age at arrival in the U.S.. In odd columns I report the results of a linear regression model, whereas even columns display the results of a model using categorical variables for intervals of age at arrival in the U.S.. In each specification I control for a sex indicator, a dummy indicator for the presence of both parents in the household, family fixed effects and, when specified, a birth order control. All these controls are interacted with a dummy variable indicating the immigration status. The sample in each column consists of native and immigrant siblings. In the immigrant sample of the first four columns I keep individuals arrived in the U.S. at the age of 17 or younger, with at least one sibling in the sample. In columns five to eight the immigrant sample consists of immigrant individuals arrived in the U.S. at the age of 17 or younger, with at least one sibling in the sample, with the additional requirement that each sibling arrived in the U.S. in the same calendar year. Heteroskedasticity robust standard errors are clustered at the family level and they are shown in parenthesis. *, ** , and *** denote significance at the 10%, 5%, 1% level, respectively.

Table A.6: The effect of age at arrival on mathematics knowledge; family fixed effects regressions controlling for presence of both parents

Sample	Full				Same Year of Arrival			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Age At Arrival	-0.04*		-0.03		-0.03		-0.01	
	(0.02)		(0.04)		(0.03)		(0.04)	
Age At Arrival 6-8		-0.23		-0.23		-0.27		-0.25
		(0.17)		(0.18)		(0.19)		(0.20)
Age At Arrival 9-11		-0.28		-0.25		-0.20		-0.14
		(0.19)		(0.21)		(0.24)		(0.27)
Age At Arrival 12-14		-0.50**		-0.50*		-0.43		-0.40
		(0.23)		(0.30)		(0.27)		(0.35)
Age At Arrival 15-17		-0.44		-0.41		-0.38		-0.31
		(0.27)		(0.36)		(0.30)		(0.42)
Family F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Both Parents	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Birth Order	No	No	Yes	Yes	No	No	Yes	Yes
N	4511	4511	4511	4511	4476	4476	4476	4476

The Table above reports the results of regressions of mathematics knowledge on immigrants' age at arrival in the U.S.. In odd columns I report the results of a linear regression model, whereas even columns display the results of a model using categorical variables for intervals of age at arrival in the U.S.. In each specification I control for a sex indicator, a dummy indicator for the presence of both parents in the household, family fixed effects and, when specified, a birth order control. All these controls are interacted with a dummy variable indicating the immigration status. The sample in each column consists of native and immigrant siblings. In the immigrant sample of the first four columns I keep individuals arrived in the U.S. at the age of 17 or younger, with at least one sibling in the sample. In columns five to eight the immigrant sample consists of immigrant individuals arrived in the U.S. at the age of 17 or younger, with at least one sibling in the sample, with the additional requirement that each sibling arrived in the U.S. in the same calendar year. Heteroskedasticity robust standard errors are clustered at the family level and they are shown in parenthesis. *, ** , and *** denote significance at the 10%, 5%, 1% level, respectively.

Table A.7: The effect of age at arrival on illicit activities; family fixed effects regressions controlling for presence of both parents

Sample	Full				Same Year of Arrival			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Age At Arrival	-0.06		-0.09		-0.08		-0.16	
	(0.06)		(0.08)		(0.07)		(0.11)	
Age At Arrival 6-8		-0.50		-0.55		-0.56		-0.66
		(0.37)		(0.38)		(0.47)		(0.48)
Age At Arrival 9-11		-0.74		-0.84*		-0.98		-1.23*
		(0.48)		(0.50)		(0.62)		(0.66)
Age At Arrival 12-14		-0.56		-0.75		-0.91		-1.33**
		(0.46)		(0.51)		(0.56)		(0.64)
Age At Arrival 15-17		-0.66		-0.93*		-1.00*		-1.59**
		(0.47)		(0.55)		(0.57)		(0.70)
Family F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Both Parents	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Birth Order	No	No	Yes	Yes	No	No	Yes	Yes
N	4158	4158	4158	4158	4126	4126	4126	4126

The Table above reports the results of regressions of illicit behaviours on immigrants' age at arrival in the U.S.. In odd columns I report the results of a linear regression model, whereas even columns display the results of a model using categorical variables for intervals of age at arrival in the U.S.. In each specification I control for a sex indicator, a dummy indicator for the presence of both parents in the household, family fixed effects and, when specified, a birth order control. All these controls are interacted with a dummy variable indicating the immigration status. The sample in each column consists of native and immigrant siblings. In the immigrant sample of the first four columns I keep individuals arrived in the U.S. at the age of 17 or younger, with at least one sibling in the sample. In columns five to eight the immigrant sample consists of immigrant individuals arrived in the U.S. at the age of 17 or younger, with at least one sibling in the sample, with the additional requirement that each sibling arrived in the U.S. in the same calendar year. Heteroskedasticity robust standard errors are clustered at the family level and they are shown in parenthesis. *, ** , and *** denote significance at the 10%, 5%, 1% level, respectively.

Excluding immigrants arriving as dropouts

Table A.8: The effect of age at arrival on years of education; family fixed effects regressions excluding dropout immigrants

Sample	Full				Same Year of Arrival			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Age At Arrival	-0.13*		-0.13		-0.15*		-0.18	
	(0.07)		(0.09)		(0.09)		(0.14)	
Age At Arrival 6-8		0.10		0.12		-0.34		-0.38
		(0.45)		(0.50)		(0.47)		(0.53)
Age At Arrival 9-11		-0.90*		-0.87		-1.38**		-1.45*
		(0.50)		(0.55)		(0.65)		(0.74)
Age At Arrival 12-14		-1.40**		-1.32		-1.82**		-1.98*
		(0.59)		(0.81)		(0.73)		(1.02)
Age At Arrival 15-17		-1.26*		-1.15		-1.59*		-1.80
		(0.70)		(1.01)		(0.83)		(1.23)
Family F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Birth Order	No	No	Yes	Yes	No	No	Yes	Yes
N	232	232	232	232	191	191	191	191

The Table above reports the results of regressions of years of education on immigrants' age at arrival in the U.S.. In odd columns I report the results of a linear regression model, whereas even columns display the results of a model using categorical variables for intervals of age at arrival in the U.S.. In each specification I control for a sex indicator, family fixed effects and, when specified, a birth order control. In the first four columns the sample consists of immigrant individuals arrived in the U.S. at the age of 17 or younger, with at least one sibling in the sample. In columns five to eight the sample consists of immigrant individuals arrived in the U.S. at the age of 17 or younger, with at least one sibling in the sample, with the additional requirement that each sibling arrived in the U.S. in the same calendar year. I exclude from the sample immigrant individuals that do not attend U.S. schools. Heteroskedasticity robust standard errors are clustered at the family level and they are shown in parenthesis. *, ** , and *** denote significance at the 10%, 5%, 1% level, respectively.

Controlling for time trends in educational system

Table A.9: The effect of age at arrival on years of education; controlling for time trends

Sample	Full				Same Year of Arrival			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Age At Arrival	-0.16**		-0.13		-0.18**		-0.19	
	(0.07)		(0.09)		(0.08)		(0.13)	
Age At Arrival 6-8		0.05		0.11		-0.38		-0.38
		(0.44)		(0.48)		(0.45)		(0.51)
Age At Arrival 9-11		-0.98**		-0.88*		-1.47**		-1.46**
		(0.47)		(0.52)		(0.61)		(0.70)
Age At Arrival 12-14		-1.48**		-1.29*		-1.91***		-1.94**
		(0.58)		(0.77)		(0.71)		(0.96)
Age At Arrival 15-17		-1.84**		-1.56		-2.25***		-2.26*
		(0.73)		(0.97)		(0.83)		(1.17)
Family F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Birth Order	No	No	Yes	Yes	No	No	Yes	Yes
N	4747	4747	4747	4747	4704	4704	4704	4704

The Table above reports the results of regressions of years of education on immigrants' age at arrival in the U.S.. In odd columns I report the results of a linear regression model, whereas even columns display the results of a model using categorical variables for intervals of age at arrival in the U.S.. In each specification I control for a sex indicator, family fixed effects and, when specified, a birth order control, each interacted with immigration status. In the first four columns the immigrant sample consists of individuals arrived in the U.S. at the age of 17 or younger, with at least one sibling in the sample. In columns five to eight the immigrant sample consists of immigrant individuals arrived in the U.S. at the age of 17 or younger, with at least one sibling in the sample, with the additional requirement that each sibling arrived in the U.S. in the same calendar year. Heteroskedasticity robust standard errors are clustered at the family level and they are shown in parenthesis. *, ** , and *** denote significance at the 10%, 5%, 1% level, respectively.

Excluding immigrants arrived after the age of 14

Table A.10: The Effect of Age at Arrival on Years of Education; Family Fixed Effects Regressions on Small Sample

Sample	Full		Same Year of Arrival	
	(1)	(2)	(3)	(4)
Age At Arrival	-0.16*	-0.16	-0.21**	-0.28
	(0.08)	(0.11)	(0.10)	(0.17)
Family F.E.	Yes	Yes	Yes	Yes
Birth Order	No	Yes	No	Yes
N	186	186	152	152

The Table above reports the results of regressions of years of education on immigrants' age at arrival in the U.S.. In each specification I control for a sex indicator, family fixed effects and, when specified, a birth order control. In the first two columns the sample consists of immigrant individuals arrived in the U.S. at the age of 13 or younger, with at least one sibling in the sample. In columns three and four the sample consists of immigrant individuals arrived in the U.S. at the age of 13 or younger, with at least one sibling in the sample, with the additional requirement that each sibling arrived in the U.S. in the same calendar year. Heteroskedasticity robust standard errors are clustered at the family level and they are shown in parenthesis. *, ** , and *** denote significance at the 10%, 5%, 1% level, respectively.

Table A.11: The Effect of age at arrival on various English outcomes; family fixed effects regressions on small sample

Outcome	Word Knowledge				Paragraph Comprehension			
Sample	Full		Same Year of Arrival		Full		Same Year of Arrival	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Age At Arrival	-0.06**	-0.05	-0.06*	-0.04	-0.06*	-0.09*	-0.07*	-0.13
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.05)	(0.04)	(0.08)
Family F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Birth Order	No	Yes	No	Yes	No	Yes	No	Yes
N	4447	4447	4417	4417	4447	4447	4417	4417

The Table above reports the results of regressions of English knowledge on immigrants' age at arrival in the U.S.. In each specification I control for a sex indicator, family fixed effects and, when specified, a birth order control. All these controls are interacted with a dummy variable indicating the immigration status. The sample in each column consists of native and immigrant siblings. In the immigrant sample of the first two columns I keep individuals arrived in the U.S. at the age of 13 or younger, with at least one sibling in the sample. In columns three and four the immigrant sample consists of immigrant individuals arrived in the U.S. at the age of 13 or younger, with at least one sibling in the sample, with the additional requirement that each sibling arrived in the U.S. in the same calendar year. Heteroskedasticity robust standard errors are clustered at the family level and they are shown in parenthesis. *, ** , and *** denote significance at the 10%, 5%, 1% level, respectively.

Table A.12: The effect of age at arrival on various mathematics outcomes; family fixed effects regressions on small sample

Outcome	Arithmetic Knowledge				Numerical Operations				Mathematics Knowledge			
Sample	Full		Same Year of Arrival		Full		Same Year of Arrival		Full		Same Year of Arrival	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Age At Arrival	-0.01	-0.07*	0.00	-0.08	-0.05	-0.05	-0.07	-0.09	-0.04	-0.05	-0.03	-0.03
	(0.04)	(0.04)	(0.05)	(0.06)	(0.05)	(0.06)	(0.05)	(0.07)	(0.03)	(0.04)	(0.03)	(0.05)
Family F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Birth Order	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
N	4447	4447	4417	4417	4447	4447	4417	4417	4447	4447	4417	4417

The Table above reports the results of regressions of mathematics knowledge on immigrants' age at arrival in the U.S.. In each specification I control for a sex indicator, family fixed effects and, when specified, a birth order control. All these controls are interacted with a dummy variable indicating the immigration status. The sample in each column consists of native and immigrant siblings. In the immigrant sample of the first two columns I keep individuals arrived in the U.S. at the age of 13 or younger, with at least one sibling in the sample. In columns three and four the immigrant sample consists of immigrant individuals arrived in the U.S. at the age of 13 or younger, with at least one sibling in the sample, with the additional requirement that each sibling arrived in the U.S. in the same calendar year. Heteroskedasticity robust standard errors are clustered at the family level and they are shown in parenthesis. *, ** , and *** denote significance at the 10%, 5%, 1% level, respectively.

Table A.13: The effect of age at arrival on illicit behaviours; family fixed effects regressions on small sample

Sample	Full		Same Year of Arrival	
	(1)	(2)	(3)	(4)
Age At Arrival	-0.08	-0.12	-0.12	-0.23
	(0.09)	(0.12)	(0.10)	(0.15)
Family F.E.	Yes	Yes	Yes	Yes
Birth Order	No	Yes	No	Yes
N	4102	4102	4075	4075

The Table above reports the results of regressions of illicit behaviours on immigrants' age at arrival in the U.S.. In each specification I control for a sex indicator, family fixed effects and, when specified, a birth order control. All these controls are interacted with a dummy variable indicating the immigration status. The sample in each column consists of native and immigrant siblings. In the immigrant sample of the first two columns I keep individuals arrived in the U.S. at the age of 13 or younger, with at least one sibling in the sample. In columns three and four the immigrant sample consists of immigrant individuals arrived in the U.S. at the age of 13 or younger, with at least one sibling in the sample, with the additional requirement that each sibling arrived in the U.S. in the same calendar year. Heteroskedasticity robust standard errors are clustered at the family level and they are shown in parenthesis. *, ** , and *** denote significance at the 10%, 5%, 1% level, respectively.

Heterogeneity

Table A.14: The effect of age at arrival on schooling; heterogeneity based on Mexican origin

Sample	Full		Same Year of Arrival	
	(1)	(2)	(3)	(4)
Age At Arrival	-0.12		-0.17	
	(0.11)		(0.15)	
Age At Arrival*Mexican	-0.08		-0.09	
	(0.15)		(0.19)	
Age At Arrival 6-8		-0.23		-0.75
		(0.47)		(0.50)
Age At Arrival 9-11		-0.93		-1.45*
		(0.63)		(0.79)
Age At Arrival 12-14		-1.36		-2.04*
		(0.90)		(1.08)
Age At Arrival 15-17		-1.21		-1.87
		(1.12)		(1.32)
Age At Arrival 6-8*Mexican		1.60		1.74
		(1.22)		(1.09)
Age At Arrival 9-11*Mexican		0.51		0.40
		(0.77)		(1.08)

Age At Arrival 12-14*Mexican		0.70		0.98
		(1.02)		(1.37)
Age At Arrival 15-17*Mexican		-0.56		-0.47
		(1.31)		(1.49)
Family F.E.	Yes	Yes	Yes	Yes
Birth Order	Yes	Yes	Yes	Yes
N	256	256	213	213

The Table above reports the results of regressions of schooling on immigrants' age at arrival in the U.S.. In odd columns I report the results of a linear regression model, whereas even columns display the results of a model using categorical variables for intervals of age at arrival in the U.S.. In each specification I control for a sex indicator, family fixed effects and a birth order control. All these controls are interacted with a dummy variable indicating the immigration status. In the immigrant sample I keep individuals arrived in the U.S. at the age of 17 or younger, with at least one sibling in the sample. Heteroskedasticity robust standard errors are clustered at the family level and they are shown in parenthesis. *, ** , and *** denote significance at the 10%, 5%, 1% level, respectively.

Table A.15: The effect of age at arrival on English knowledge; heterogeneity based on Mexican origin

Outcome	Word Knowledge				Paragraph Comprehension			
Sample	Full		Same Year of Arrival		Full		Same Year of Arrival	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Age At Arrival	-0.04		-0.04		-0.08		-0.12*	
	(0.03)		(0.03)		(0.05)		(0.07)	
Age At Arrival*Mexican	-0.03		0.02		-0.04		-0.01	
	(0.05)		(0.06)		(0.05)		(0.06)	
Age At Arrival 6-8		0.19		0.04		0.02		-0.15
		(0.15)		(0.15)		(0.20)		(0.20)
Age At Arrival 9-11		-0.07		-0.15		-0.32		-0.42
		(0.13)		(0.14)		(0.22)		(0.28)
Age At Arrival 12-14		-0.36		-0.54**		-1.03***		-1.25***
		(0.26)		(0.25)		(0.31)		(0.37)
Age At Arrival 15-17		-0.37		-0.45		-1.23***		-1.53***
		(0.33)		(0.34)		(0.45)		(0.49)
Age At Arrival 6-8*Mexican		-1.11***		-0.40		0.16		0.67***
		(0.37)		(0.37)		(0.37)		(0.25)

Age At Arrival 9-11*Mexican		-0.65		-0.03		0.16		0.51*
		(0.42)		(0.41)		(0.40)		(0.30)
Age At Arrival 12-14*Mexican		-0.70		0.14		0.18		0.72
		(0.52)		(0.53)		(0.47)		(0.48)
Age At Arrival 15-17*Mexican		-0.77		-0.12		0.06		0.60
		(0.61)		(0.58)		(0.58)		(0.55)
Family F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Birth Order	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	4511	4511	4476	4476	4511	4511	4476	4476

The Table above reports the results of regressions of several outcomes on immigrants' age at arrival in the U.S.. In odd columns I report the results of a linear regression model, whereas even columns display the results of a model using categorical variables for intervals of age at arrival in the U.S.. In each specification I control for a sex indicator, family fixed effects and a birth order control. In the immigrant sample I keep individuals arrived in the U.S. at the age of 17 or younger, with at least one sibling in the sample. Heteroskedasticity robust standard errors are clustered at the family level and they are shown in parenthesis. *, ** , and *** denote significance at the 10%, 5%, 1% level, respectively.

Table A.16: The effect of age at arrival on mathematics knowledge; heterogeneity based on Mexican origin

Outcome	Arithmetic				Num. Operations				Maths Knowledge			
Sample	Full		Same Year of A.		Full		Same Year of A.		Full		Same Year of A.	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
A. At A.	-0.03		-0.03		-0.01		-0.03		-0.03		-0.00	
	(0.04)		(0.05)		(0.06)		(0.07)		(0.04)		(0.05)	
A. At A.*Mex.	-0.07		-0.08		-0.09		-0.08		0.00		0.01	
	(0.05)		(0.06)		(0.07)		(0.08)		(0.04)		(0.04)	
A. At A. 6-8		-0.28*		-0.29*		0.00		-0.25		-0.16		-0.24
		(0.15)		(0.17)		(0.33)		(0.34)		(0.19)		(0.21)
A. At A. 9-11		-0.13		-0.11		0.01		-0.27		-0.20		-0.09
		(0.23)		(0.29)		(0.37)		(0.44)		(0.24)		(0.29)
A. At A. 12-14		-0.41		-0.47		-0.43		-0.82		-0.47		-0.40
		(0.27)		(0.32)		(0.48)		(0.58)		(0.35)		(0.39)
A. At A. 15-17		-0.19		-0.31		0.08		-0.10		-0.17		-0.08
		(0.38)		(0.45)		(0.75)		(0.82)		(0.42)		(0.47)
A. At A. 6-8*Mex.		0.14		0.19		-0.47		0.26		-0.51*		-0.26
		(0.25)		(0.24)		(0.63)		(0.73)		(0.29)		(0.32)

A. At A. 9-11*Mex.		-0.40		-0.47		-0.50		0.30		-0.38		-0.36
		(0.29)		(0.31)		(0.69)		(0.83)		(0.27)		(0.35)
A. At A. 12-14*Mex.		-0.45		-0.50		-0.48		0.24		-0.23		-0.09
		(0.37)		(0.46)		(0.75)		(0.91)		(0.41)		(0.49)
A. At A. 15-17*Mex.		-1.05**		-1.07**		-1.39		-0.91		-0.68*		-0.59
		(0.42)		(0.48)		(0.94)		(1.08)		(0.40)		(0.47)
Family F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Birth Order	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	4511	4511	4476	4476	4511	4511	4476	4476	4511	4511	4476	4476

The Table above reports the results of regressions of several outcomes on immigrants' age at arrival in the U.S.. In odd columns I report the results of a linear regression model, whereas even columns display the results of a model using categorical variables for intervals of age at arrival in the U.S.. In each specification I control for a sex indicator, family fixed effects and a birth order control. In the immigrant sample I keep individuals arrived in the U.S. at the age of 17 or younger, with at least one sibling in the sample. Heteroskedasticity robust standard errors are clustered at the family level and they are shown in parenthesis. *, ** , and *** denote significance at the 10%, 5%, 1% level, respectively.

Table A.17: The effect of age at arrival on illicit behaviors; heterogeneity based on Mexican origin

Sample	Full		Same Year of Arrival	
	(1)	(2)	(3)	(4)
Age At Arrival	-0.11		-0.18	
	(0.11)		(0.13)	
Age At Arrival*Mexican	0.06		0.10	
	(0.08)		(0.09)	
Age At Arrival 6-8		-0.65		-0.74
		(0.44)		(0.53)
Age At Arrival 9-11		-0.96		-1.40**
		(0.60)		(0.70)
Age At Arrival 12-14		-0.82		-1.47**
		(0.58)		(0.66)
Age At Arrival 15-17		-0.93		-1.65**
		(0.63)		(0.73)
Age At Arrival 6-8*Mexican		0.56		0.95*
		(0.49)		(0.54)
Age At Arrival 9-11*Mexican		0.62		1.36**
		(0.61)		(0.68)

Age At Arrival 12-14*Mexican		0.44		1.26**
		(0.57)		(0.61)
Age At Arrival 15-17*Mexican		0.23		1.03*
		(0.58)		(0.63)
Family F.E.	Yes	Yes	Yes	Yes
Birth Order	Yes	Yes	Yes	Yes
N	4158	4158	4126	4126

The Table above reports the results of regressions of illicit behaviors outcomes on immigrants' age at arrival in the U.S.. In odd columns I report the results of a linear regression model, whereas even columns display the results of a model using categorical variables for intervals of age at arrival in the U.S.. In each specification I control for a sex indicator, family fixed effects and a birth order control. In the immigrant sample I keep individuals arrived in the U.S. at the age of 17 or younger, with at least one sibling in the sample. Heteroskedasticity robust standard errors are clustered at the family level and they are shown in parenthesis. *, ** , and *** denote significance at the 10%, 5%, 1% level, respectively.

Table A.18: The effect of age at arrival on schooling; heterogeneity based on arrival from English speaking country

Sample	Full		Same Year of Arrival	
	(1)	(2)	(3)	(4)
Age At Arrival	-0.11		-0.14	
	(0.08)		(0.11)	
Age At Arrival*English	-0.17		-0.28	
	(0.19)		(0.18)	
Age At Arrival 6-8		0.41		0.05
		(0.50)		(0.51)
Age At Arrival 9-11		-0.57		-0.97
		(0.49)		(0.70)
Age At Arrival 12-14		-0.78		-1.23
		(0.66)		(0.91)
Age At Arrival 15-17		-1.30		-1.78
		(0.89)		(1.12)
Age At Arrival 6-8*English		-1.00		-1.86
		(1.22)		(1.19)
Age At Arrival 9-11*English		-0.71		-1.41
		(1.00)		(0.87)

Age At Arrival 12-14*English		-1.86**		-2.24***
		(0.83)		(0.82)
Age At Arrival 15-17*English		0.52		0.03
		(0.99)		(0.96)
Family F.E.	Yes	Yes	Yes	Yes
Birth Order	Yes	Yes	Yes	Yes
N	256	256	213	213

The Table above reports the results of regressions of schooling on immigrants' age at arrival in the U.S.. In odd columns I report the results of a linear regression model, whereas even columns display the results of a model using categorical variables for intervals of age at arrival in the U.S.. In each specification I control for a sex indicator, family fixed effects and a birth order control. In the immigrant sample I keep individuals arrived in the U.S. at the age of 17 or younger, with at least one sibling in the sample. Heteroskedasticity robust standard errors are clustered at the family level and they are shown in parenthesis. *, ** , and *** denote significance at the 10%, 5%, 1% level, respectively.

Table A.19: The effect of age at arrival on English knowledge; heterogeneity based on arrival from English speaking country

Outcome	Word Knowledge				Paragraph Comprehension			
Sample	Full		Same Year of Arrival		Full		Same Year of Arrival	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Age At Arrival	-0.06*		-0.03		-0.08*		-0.11*	
	(0.03)		(0.03)		(0.05)		(0.06)	
Age At Arrival*English	0.02		-0.02		-0.04		-0.05	
	(0.04)		(0.03)		(0.05)		(0.05)	
Age At Arrival 6-8		-0.09		-0.02		0.01		-0.03
		(0.18)		(0.17)		(0.19)		(0.22)
Age At Arrival 9-11		-0.22		-0.10		-0.18		-0.28
		(0.19)		(0.19)		(0.28)		(0.34)
Age At Arrival 12-14		-0.54*		-0.44		-0.82**		-0.97**
		(0.30)		(0.30)		(0.35)		(0.43)
Age At Arrival 15-17		-0.61		-0.45		-1.08**		-1.30***
		(0.37)		(0.39)		(0.43)		(0.50)
Age At Arrival 6-8*English		0.37		-0.12		0.37		0.12
		(0.37)		(0.32)		(0.38)		(0.37)

Age At Arrival 9-11*English		0.25		-0.03		-0.08		-0.02
		(0.20)		(0.16)		(0.28)		(0.36)
Age At Arrival 12-14*English		0.12		-0.17		-0.62**		-0.61*
		(0.27)		(0.25)		(0.30)		(0.37)
Age At Arrival 15-17*English		0.19		-0.17		-0.06		0.01
		(0.31)		(0.31)		(0.35)		(0.43)
Family F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Birth Order	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	4511	4511	4476	4476	4511	4511	4476	4476

The Table above reports the results of regressions of several outcomes on immigrants' age at arrival in the U.S.. In odd columns I report the results of a linear regression model, whereas even columns display the results of a model using categorical variables for intervals of age at arrival in the U.S.. In each specification I control for a sex indicator, family fixed effects and a birth order control. In the immigrant sample I keep individuals arrived in the U.S. at the age of 17 or younger, with at least one sibling in the sample. Heteroskedasticity robust standard errors are clustered at the family level and they are shown in parenthesis. *, ** , and *** denote significance at the 10%, 5%, 1% level, respectively.

Table A.20: The effect of age at arrival on mathematics knowledge; heterogeneity based on arrival from English speaking country

Outcome	Arithmetic				Num. Operations				Maths Knowledge			
Sample	Full		Same Year of A.		Full		Same Year of A.		Full		Same Year of A.	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
A. At A.	-0.06		-0.07		-0.08*		-0.09		-0.04		-0.02	
	(0.04)		(0.05)		(0.05)		(0.06)		(0.04)		(0.05)	
A. At A.*Eng.	0.04		0.04		0.18***		0.13*		0.04		0.06	
	(0.04)		(0.05)		(0.06)		(0.07)		(0.05)		(0.05)	
A. At A. 6-8		-0.16		-0.14		-0.50**		-0.61**		-0.35*		-0.31
		(0.15)		(0.18)		(0.24)		(0.25)		(0.19)		(0.20)
A. At A. 9-11		-0.42*		-0.39		-0.51*		-0.60*		-0.42**		-0.35
		(0.25)		(0.30)		(0.28)		(0.32)		(0.18)		(0.22)
A. At A. 12-14		-0.61*		-0.64		-1.12***		-1.37***		-0.74***		-0.66**
		(0.32)		(0.39)		(0.36)		(0.41)		(0.27)		(0.31)
A. At A. 15-17		-0.68		-0.78		-1.02*		-1.18**		-0.62*		-0.53
		(0.43)		(0.53)		(0.57)		(0.60)		(0.34)		(0.38)
A. At A. 6-8*Eng.		-0.05		0.19		1.91***		1.80***		0.38		0.08
		(0.46)		(0.61)		(0.51)		(0.61)		(0.51)		(0.52)

A. At A. 9-11*Eng.		0.54		0.74		1.15***		0.98**		0.34		0.41
		(0.48)		(0.64)		(0.32)		(0.41)		(0.36)		(0.55)
A. At A. 12-14*Eng.		-0.21		-0.11		1.63***		1.60***		0.78***		0.75**
		(0.31)		(0.38)		(0.33)		(0.39)		(0.26)		(0.34)
A. At A. 15-17*Eng.		0.35		0.54		1.80***		1.60***		0.41		0.42
		(0.73)		(0.84)		(0.48)		(0.50)		(0.35)		(0.50)
Family F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Birth Order	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	4511	4511	4476	4476	4511	4511	4476	4476	4511	4511	4476	4476

The Table above reports the results of regressions of several outcomes on immigrants' age at arrival in the U.S.. In odd columns I report the results of a linear regression model, whereas even columns display the results of a model using categorical variables for intervals of age at arrival in the U.S.. In each specification I control for a sex indicator, family fixed effects and a birth order control. In the immigrant sample I keep individuals arrived in the U.S. at the age of 17 or younger, with at least one sibling in the sample. Heteroskedasticity robust standard errors are clustered at the family level and they are shown in parenthesis. *, ** , and *** denote significance at the 10%, 5%, 1% level, respectively.

Table A.21: The effect of age at arrival on illicit behaviors; heterogeneity based on arrival from English speaking country

Sample	Full		Same Year of Arrival	
	(1)	(2)	(3)	(4)
Age At Arrival	-0.09		-0.16	
	(0.10)		(0.13)	
Age At Arrival*English	-0.00		0.00	
	(0.09)		(0.10)	
Age At Arrival 6-8		-0.59		-0.71
		(0.49)		(0.58)
Age At Arrival 9-11		-1.19		-1.62*
		(0.77)		(0.88)
Age At Arrival 12-14		-1.00		-1.67*
		(0.78)		(0.89)
Age At Arrival 15-17		-1.20		-1.94**
		(0.81)		(0.94)
Age At Arrival 6-8*English		0.13		0.12
		(0.54)		(0.67)
Age At Arrival 9-11*English		0.81		1.03
		(0.76)		(0.86)

Age At Arrival 12-14*English		0.17		0.58
		(0.74)		(0.84)
Age At Arrival 15-17*English		0.52		0.90
		(0.85)		(0.93)
Family F.E.	Yes	Yes	Yes	Yes
Birth Order	Yes	Yes	Yes	Yes
N	4158	4158	4126	4126

The Table above reports the results of regressions of illicit behaviors outcomes on immigrants' age at arrival in the U.S.. In odd columns I report the results of a linear regression model, whereas even columns display the results of a model using categorical variables for intervals of age at arrival in the U.S.. In each specification I control for a sex indicator, family fixed effects and a birth order control. In the immigrant sample I keep individuals arrived in the U.S. at the age of 17 or younger, with at least one sibling in the sample. Heteroskedasticity robust standard errors are clustered at the family level and they are shown in parenthesis. *, ** , and *** denote significance at the 10%, 5%, 1% level, respectively.

Accounting exercise

Table A.22: The mincerian regressions

Sample	Immigrants				Natives			
Dependent Variable	P(Working)	P(Full Time)	H. Wage	Log Earn	P(Working)	P(Full Time)	H. Wage	Log Earn
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Education	0.01**	0.00	0.06***	0.07***	0.00**	0.01***	0.06***	0.07***
	(0.00)	(0.01)	(0.01)	(0.02)	(0.00)	(0.00)	(0.00)	(0.00)
Word. Know.	-0.03***	-0.04***	-0.08***	-0.09**	0.02***	-0.01*	0.03***	0.04***
	(0.01)	(0.01)	(0.02)	(0.04)	(0.00)	(0.00)	(0.01)	(0.01)
Par. Compreh.	0.04***	0.04***	0.10***	0.15***	0.01***	0.02***	0.02***	0.06***
	(0.01)	(0.01)	(0.02)	(0.04)	(0.00)	(0.00)	(0.01)	(0.01)
Arit. Reason.	0.04***	0.00	0.04	0.00	-0.01***	-0.00	0.02***	0.03**
	(0.01)	(0.02)	(0.03)	(0.06)	(0.00)	(0.00)	(0.01)	(0.01)
Num. Oper.	0.03***	0.06***	0.13***	0.23***	0.03***	0.03***	0.10***	0.17***
	(0.01)	(0.01)	(0.02)	(0.04)	(0.00)	(0.00)	(0.00)	(0.01)
Math. Know.	-0.07***	-0.04**	-0.00	-0.09	-0.01**	0.02***	0.07***	0.07***
	(0.01)	(0.02)	(0.02)	(0.06)	(0.00)	(0.00)	(0.01)	(0.01)
Illicit	-0.01**	-0.01	-0.03*	-0.03	-0.01***	-0.04***	-0.01***	-0.06***
	(0.01)	(0.01)	(0.02)	(0.04)	(0.00)	(0.00)	(0.00)	(0.01)
N	6177	5374	4357	4479	116148	100772	83048	84999

The Table above reports the results of regressions of several labor market outcomes on education, different cognitive skills and the measure of illicit behaviours. In each column I also control for two ethnicity indicators, gender, year fixed effects, potential experience and its square. Heteroskedasticity robust standard errors are shown in parenthesis. *, ** , and *** denote significance at the 10%, 5%, 1% level, respectively.

Appendix B

Appendix to chapter 2: The Immigrant American Dream

Samples and countries of origin

Table B.1: Sample selection; NLSY79 and NLSY79 Children And Young Adults

	Number of Individuals
NLSY79	
Original Dataset	12,686
Discard Individuals Arrived Before 18	12,548
Discard Individuals With Unknown Age At Arrival	12,548
Discard 1st Gen. Immigrants With Father or Mother or Paternal Grandfather Born in U.S.	12,354
Discard 2nd Gen. Immigrants From Father's Side, With Paternal Grandfather Born in U.S.	12,325
NLSY79 Children And Young Adults	
Original Dataset	11,521
Discard Children Never Interviewed	10,503
Discard Children Born Abroad	10,499
Discard Children of 1st Gen. Mother Arrived Before 18	10,402
Discard Children of 1st Gen. Mother With Unknown Age At Arrival	10,402
Discard Children of 1st Gen. Mother With Father or Mother or Paternal Grandfather Born in U.S.	10,244
Discard Children of 2nd Gen. Mother From Father's Side, With Paternal Grandfather Born in U.S.	10,223

Table B.2: Countries of origin; NLSY79

Country of Birth	1st Generation	2nd Generation		3rd Generation
		Mother	Father	
Mexico	231	256	214	177
Cuba	64	27	36	4
Dominican Republic	22	3	5	
Ecuador	16	2	1	
Canada	15	52	34	51
Jamaica	15	4	9	5
Portugal	15	7	8	16
Hong Kong	11			
Italy	10	13	29	190
Philippines	10	5	16	6
Colombia	7	4	1	
Guatemala	7	1	1	
Poland	6	10	10	68
El Salvador	5			
Haiti	5	2		
India	5		3	1
Iran	5			
Chile	4	2		
Netherlands	4	3	5	12
Peru	4			
Venezuela	4		1	1
Yugoslavia	4	8	9	11
Argentina	3	3	1	

Barbados	3		1	1
England	3	15	7	26
Greece	3	3	5	6
Honduras	3			1
Nicaragua	3	1		
Nigeria	3			
Panama	3			
Costa Rica	2			
Germany	2	31	17	121
Guyana	2			
Iraq	2			
Japan	2	14	1	3
South Korea	2			1
Trinidad & Tobago	2	2	2	1
Bahamas	1			3
Belgium	1	1		3
Brazil	1	2	1	
Cambodia	1			
French Guiana	1			
Guinea-Bisseau	1			
Israel	1	3		
Libya	1			
Switzerland	1			4
Thailand	1			
Togo	1			
Turkey	1			2
Uruguay	1			

Vietnam	1			
Virgin Islands	1			
France		7	3	7
Ireland		7	3	36
Scotland		7	5	15
Denmark		6	1	6
Hungary		4	7	12
China		3	4	2
Panama		3		
Czechoslovakia		2	5	18
Finland		2		8
Norway		2	1	13
Africa, n.s.		1	1	
Austria		1	1	16
Cyprus		1	1	
Egypt		1		
Jordan		1	1	
Iceland		1		1
Liechtenstein		1	1	2
Libya		1	2	
Malta		1		
New Zealand		1		
Romania		1	3	
South Africa		1	1	1
Spain		1	3	14
Wales		1		2
Australia			2	2

Lebanon			2	3
Luxembourg			1	
Peru			1	
Caribbean				2
Europe, n.s.				4
Luxembourg				1
Saudi Arabi				3
U.S.S.R.				4
Not U.S., n.s.	15	25	27	87
U.S.		197	263	

Robustness checks

Different empirical specification: outcomes in levels and family fixed effects

Table B.3: Different empirical specification: outcomes in levels and family fixed effects

	Y. of Sch.		P(High School)		P(College)		Maths		Read. Comp.		Word Reco.	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(9)	(7)	(10)	(11)	(12)
Constant	0.21**		-0.13***		0.07***		0.38***		0.18***		0.37***	
	(0.08)		(0.01)		(0.01)		(0.02)		(0.02)		(0.02)	
γ_2	1.57***		0.23***		0.06		0.29**		0.65***		0.75***	
	(0.41)		(0.05)		(0.06)		(0.11)		(0.12)		(0.12)	
γ_3	-0.19		0.05		-0.02		0.10		0.04		0.08	
	(0.28)		(0.04)		(0.05)		(0.07)		(0.07)		(0.07)	
γ_4	-0.09		0.03		-0.01		0.00		0.01		0.02	
	(0.30)		(0.02)		(0.05)		(0.06)		(0.06)		(0.06)	
Child		0.22***		-0.13***		0.07***		0.44***		0.22***		0.42***
		(0.08)		(0.01)		(0.01)		(0.02)		(0.02)		(0.02)
γ_2^{FE}		1.19***		0.19***		0.08		0.23**		0.54***		0.61***
		(0.36)		(0.04)		(0.06)		(0.10)		(0.10)		(0.11)
γ_3^{FE}		-0.23		0.05		-0.04		0.03		-0.02		0.02
		(0.28)		(0.03)		(0.04)		(0.07)		(0.07)		(0.07)
γ_4^{FE}		0.03		0.04**		0.01		-0.02		-0.00		-0.02

		(0.25)		(0.02)		(0.05)		(0.06)		(0.06)		(0.06)
N	4378	6818	7103	10363	5308	8062	31227	34921	26534	30139	31100	34790

The Table above reports results of OLS and family-F.E. regressions of the difference between a child's and his mother's educational and cognitive achievements on immigration status. In the first column of each outcome variable we use the model in difference, while in the second one we pool children and mothers observations and we include family fixed effects. γ_2 refers to second generation immigrants. γ_3 refers to third generation immigrants. γ_4 refers to fourth generation immigrants. In each column we control for gender of the individual. The sample of each column consists of the same individuals for the corresponding outcomes in the previous tables. Heteroskedasticity robust standard errors are clustered at the family level and are shown in parenthesis. *, ** , and *** denote significance at the 10%, 5%, 1% level, respectively.

Table B.4: Different empirical specification: outcomes in levels and family fixed effects

	Behavioural Problems	
	(1)	(2)
Constant	0.54***	
	(0.02)	
γ_2	0.15	
	(0.10)	
γ_3	-0.10	
	(0.08)	
γ_4	-0.24***	
	(0.08)	
Child		0.55***
		(0.02)
γ_2^{FE}		0.08
		(0.10)
γ_3^{FE}		-0.11
		(0.07)
γ_4^{FE}		-0.24***
		(0.07)
N	32787	36571

The Table above reports results of OLS and family-F.E. regressions of the difference between a child's and his mother's attitude towards being involved in problematic behaviors on immigration status. In the first column of each outcome variable we use the model in difference, while in the second one we pool children and mothers observations and we include family fixed effects. γ_2 refers to second generation immigrants. γ_3 refers to third generation immigrants. γ_4 refers to fourth generation immigrants. In each column we control for gender of the individual. The sample of each column consists of the same individuals for the corresponding outcomes in the previous tables. Heteroskedasticity robust standard errors are clustered at the family level and are shown in parenthesis. *, ** , and *** denote significance at the 10%, 5%, 1% level, respectively.

Table B.5: Different empirical specification: outcomes in levels and family fixed effects

	Both Parents	
	(1)	(2)
Constant	-0.03**	
	(0.01)	
γ_2	0.05	
	(0.06)	
γ_3	-0.02	
	(0.04)	
γ_4	0.04	
	(0.03)	
Child		-0.06***
		(0.01)
γ_2^{FE}		0.05
		(0.05)
γ_3^{FE}		-0.01
		(0.04)
γ_4^{FE}		0.04
		(0.03)
N	51076	55412

The Table above reports results of OLS and family-F.E. regressions of the difference between a child's and his mother's likelihood of living in two-parents families on immigration status. In the first column of each outcome variable we use the model in difference, while in the second one we pool children and mothers observations and we include family fixed effects. γ_2 refers to second generation immigrants. γ_3 refers to third generation immigrants. γ_4 refers to fourth generation immigrants. In each column we control for gender of the individual. The sample of each column consists of the same individuals for the corresponding outcomes in the previous tables. Heteroskedasticity robust standard errors are clustered at the family level and are shown in parenthesis. *, **, and *** denote significance at the 10%, 5%, 1% level, respectively.

Table B.6: Different empirical specification: outcomes in levels and family fixed effects

	Number Of Children		Age At First Birth		P(Teenage Mother)	
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	-1.80***		0.95***		-0.04***	
	(0.05)		(0.16)		(0.01)	
γ_2	-0.53**		0.94		0.01	
	(0.21)		(0.86)		(0.03)	
γ_3	-0.17		-0.76		-0.01	
	(0.16)		(0.69)		(0.04)	
γ_4	-0.31**		-0.01		-0.04	
	(0.14)		(0.78)		(0.03)	
Child		-1.69***		0.88***		-0.04***
		(0.04)		(0.16)		(0.01)
γ_2^{FE}		-0.51***		1.29		0.02
		(0.14)		(1.05)		(0.03)
γ_3^{FE}		-0.16		-0.75		-0.01
		(0.15)		(0.76)		(0.03)
γ_4^{FE}		-0.31**		-0.51		-0.02

		(0.13)		(0.80)		(0.02)
N	5198	7923	1955	3378	3789	6314

The Table above reports results of OLS and family-F.E. regressions of the difference between a child's and his mother's parenting outcome on immigration status. In the first column of each outcome variable we use the model in difference, while in the second one we pool chidlren and mothers observations and we include family fixed effects. γ_2 refers to second generation immigrants. γ_3 refers to third generation immigrants. γ_4 refers to fourth generation immigrants. In each column we control for gender of the individual. The sample of each column consists of the same individuals for the corresponding outcomes in the previous tables. Heteroskedasticity robust standard errors are clustered at the family level and are shown in parenthesis. *, ** , and *** denote significance at the 10%, 5%, 1% level, respectively.

Educational achievements: different sampling strategy

Table B.7: Educational achievements: different sampling strategy

	Δ Y. of Sch.			Δ P(High School)		ΔP(College)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Sample	Max Age $\geq$ 23	Max A. $\geq$ 25	Max A. $\geq$ 30	Max A. $\geq$ 18	Max A. $\geq$ 23	Max A. $\geq$ 23	Max A. $\geq$ 25
Constant	1.68***	1.34***	1.94**	-0.22***	-0.16**	0.07	0.02
	(0.34)	(0.40)	(0.81)	(0.05)	(0.07)	(0.05)	(0.06)
γ_2	1.64***	1.60***	1.92***	0.24***	0.29***	0.08	0.05
	(0.38)	(0.44)	(0.69)	(0.05)	(0.05)	(0.06)	(0.07)
γ_3	-0.34	-0.23	-0.26	0.05	0.06	-0.02	-0.01
	(0.26)	(0.29)	(0.41)	(0.04)	(0.05)	(0.05)	(0.05)
γ_4	-0.27	-0.18	-0.51	0.02	0.01	-0.03	0.00
	(0.26)	(0.30)	(0.48)	(0.02)	(0.03)	(0.05)	(0.06)
Gender	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Other Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	5288	4378	1724	7103	4386	5308	4386

The Table above reports results of OLS regressions of the difference between a child's and his mother's educational achievement on immigration status. γ_2 refers to the coefficient on the indicator for children whose mothers are first generation immigrants, that is born abroad with parents and paternal grandfather born abroad. γ_3 refers to the coefficient on the indicator for children whose mothers are second generation immigrants, that is born in the U.S. with at least one parent born abroad. γ_4 refers to the coefficient on the indicator for children whose mothers are third generation immigrants, that is born in the U.S. with parents also born in the U.S., but with the paternal grandfather born abroad. *Other Controls* refer to three ethnicity indicators, age of mother at birth and number of older siblings. Heteroskedasticity robust standard errors are clustered at the family level and are shown in parenthesis. *, ** , and *** denote significance at the 10%, 5%, 1% level, respectively.

Different measures of reading comprehension and word knowledge

Table B.8: Different measures of reading comprehension and word knowledge

	Δ Reading Comprehension		Δ Word Knowledge	
Measure	Usual	Alternative	Usual	Alternative
	(1)	(2)	(3)	(4)
Constant	1.53^{***}	1.34^{***}	0.58^{***}	0.57^{***}
	(0.08)	(0.09)	(0.08)	(0.08)
γ_2	0.48^{***}	0.55^{***}	0.59^{***}	0.53^{***}
	(0.11)	(0.13)	(0.12)	(0.12)
γ_3	0.03	0.08	0.05	0.01
	(0.08)	(0.09)	(0.07)	(0.08)
γ_4	0.10^{*}	0.15^{**}	0.10	0.08
	(0.06)	(0.06)	(0.06)	(0.06)
Gender	Yes	Yes	Yes	Yes
Other Controls	Yes	Yes	Yes	Yes
N	26534	26534	31100	31148

The Table above reports results of OLS regressions of the difference between a child's and his mother's English achievement on immigration status. γ_2 refers to the coefficient on the indicator for children whose mothers are first generation immigrants, that is born abroad with parents and paternal grandfather born abroad. γ_3 refers to the coefficient on the indicator for children whose mothers are second generation immigrants, that is born in the U.S. with at least one parent born abroad. γ_4 refers to the coefficient on the indicator for children whose mothers are third generation immigrants, that is born in the U.S. with parents also born in the U.S., but with the paternal grandfather born abroad. *Other Controls* refer to three ethnicity indicators, age of mother at birth and child's age and number of older siblings. Heteroskedasticity robust standard errors are clustered at the family level and are shown in parenthesis. *, ** , and *** denote significance at the 10%, 5%, 1% level, respectively.

Additional tables

Table B.9: Summary statistics; mothers in NLSY79 Children And Young Adults

	Native	1st Generation	2nd Generation	3rd Generation
Years of Schooling	13.47	12.95	14.03	14.05
Mathematics Knowledge	-0.16	-0.55	-0.02	0.08
English Knowledge	-0.04	-0.66	0.07	0.20
Illicit Behavior	-0.25	-0.54	-0.28	-0.22
Both Parents	0.71	0.75	0.77	0.80
Sons	2.34	2.55	2.27	2.26
Age at First Birth	24.22	24.70	26.27	26.21
White	0.79	0.45	0.71	0.89
Black	0.17	0.07	0.04	0.02
Hispanic	0.04	0.48	0.26	0.10
Individuals	3600	202	268	325

Table B.10: Principal component analysis; immigrants in NLSY79

	Whole Sample			Mothers			
Cumulative Proportion	0.73			0.60			
	Comp.1	Comp.2	Unexplained	Comp.1	Comp.2	Comp.3	Unexplained
Years of Schooling	0.50		0.19	0.50			0.21
High-School	0.30		0.72				0.71
College	0.44		0.35	0.47			0.36
Mathematics	0.49		0.27	0.49			0.29
English	0.47		0.31	0.47			0.35
Illicit Behavior		0.85	0.23			-0.68	0.43
Both Parents		-0.43	0.78			0.63	0.45
Sons	X	X	X		0.56		0.43
Age at First Birth	X	X	X		-0.55		0.34
Teenage Mother	X	X	X		0.53		0.46
Individuals	2083			684			

The Table above reports results of the principal component analysis using the sample of immigrant individuals in the NLSY79. We retain factors whose corresponding eigenvalue is larger than 1, and we report factor loadings larger than 0.3.

Table B.11: Principal component analysis; Δ child-mother immigrants in Children of NLSY79

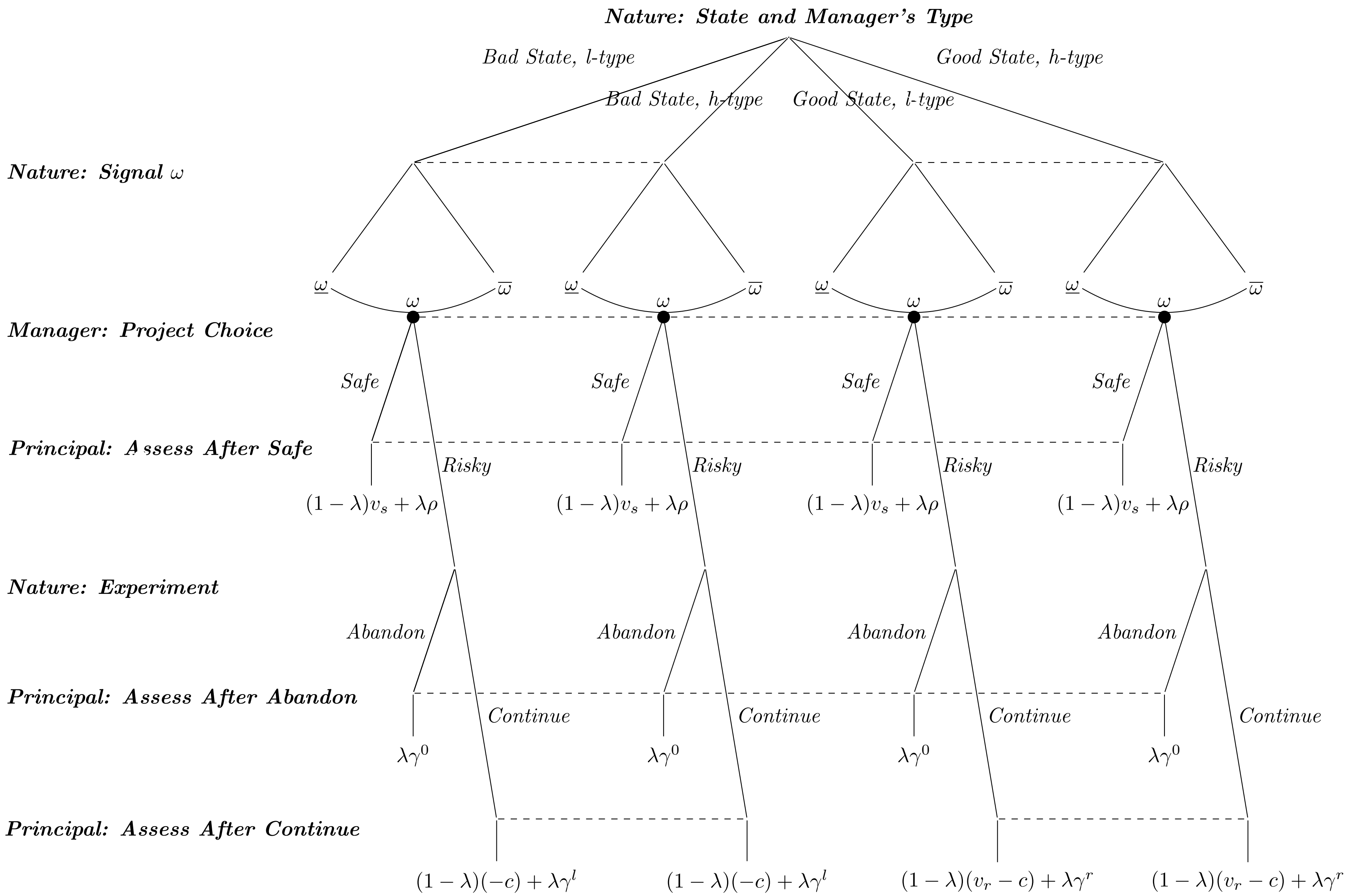
	Whole Sample				Womens				
Cumulative Proportion	0.68				0.66				
	Comp.1	Comp.2	Comp.3	Unexplained	Comp.1	Comp.2	Comp.3	Comp.4	Unexplained
Δ Years of Schooling		0.64		0.16		0.60			0.17
Δ High-School		0.35	0.42	0.50		0.51			0.45
Δ College		0.64		0.23		0.38		-0.55	0.22
Δ Mathematics	0.51			0.34	0.46				0.44
Δ English	0.60			0.12	0.60				0.14
Δ English	0.60			0.11	0.58				0.16
Δ Illicit Behavior			0.46	0.68				0.71	0.38
Δ Both Parents			0.75	0.41		0.48		0.37	0.48
Δ Sons	X	X	X	X			0.37		0.73
Δ Age at First Birth	X	X	X	X			-0.68		0.22
Δ Teenage Mother	X	X	X	X			0.61		0.37
Individuals	722				232				

The Table above reports results of the principal component analysis using the sample of immigrant individuals in the NLSY79 Children And Young Adults. The various measures are expressed as the difference between a child's and his mother's outcomes. We retain factors whose corresponding eigenvalue is larger than 1, and we report factor loadings larger than 0.3.

Appendix C

Appendix to chapter 3: A Model of Risk Taking with Experimentation and Career Concerns

Graphical game representation



Notes: The Figure above reports the timing of the economy, the information structure and the manager's payoff in every possible scenario. Dashed lines means that the Nature, the Manager or the Principal do not know what happened before. The only strategic move is the agent's project choice, denoted with a solid circle. The principal assessments follow Bayes rule.

Proofs

Beliefs

$$\gamma^r \equiv \mathbb{P}(\theta = h \,|\, x = g,\, s = g) = \frac{\mathbb{P}(\theta = h,\, x = g,\, s = g)}{\mathbb{P}(x = g,\, s = g)} =$$

$$= \frac{\int \overbrace{\mathbb{P}(\theta = h \,|\, \omega)}^{=\mathbb{P}(\theta=h)=\rho}\, \mathbb{P}(x = g,\, s = g \,|\, \omega,\, \theta = h)\, \sigma(\omega)\, dF(\omega)}{\sum_{i \in \{l,h\}} \int \underbrace{\mathbb{P}(\theta = i \,|\, \omega)}_{=\mathbb{P}(\theta=i)} \mathbb{P}(x = g,\, s = g \,|\, \omega,\, \theta = i)\, \sigma(\omega)\, dF(\omega)} =$$

$$= \rho \frac{\int \mathbb{P}(s = g \,|\, x = g,\, \omega,\, \theta = h)\, \overbrace{\mathbb{P}(x = g \,|\, \omega,\, \theta = h)}^{=\mathbb{P}(x=g \,|\, \omega)=p(\omega)}\, \sigma(\omega)\, dF(\omega)}{\sum_{i \in \{l,h\}} \mathbb{P}(\theta = i) \int \mathbb{P}(s = g \,|\, x = g,\, \omega,\, \theta = i)\, \underbrace{\mathbb{P}(x = g \,|\, \omega,\, \theta = i)}_{=\mathbb{P}(x=g \,|\, \omega)=p(\omega)}\, \sigma(\omega)\, dF(\omega)}$$

$$\gamma^0 \equiv \mathbb{P}(\theta = h \,|\, s = b) = \frac{\mathbb{P}(\theta = h,\, s = b)}{\mathbb{P}(s = b)}$$

$$= \frac{\int \overbrace{\mathbb{P}(\theta = h \,|\, \omega)}^{=\mathbb{P}(\theta=h)=\rho}\, \mathbb{P}(s = b \,|\, \omega,\, \theta = h)\, \sigma(\omega)\, dF(\omega)}{\sum_{i \in \{l,h\}} \int \underbrace{\mathbb{P}(\theta = i \,|\, \omega)}_{=\mathbb{P}(\theta=i)} \mathbb{P}(s = b \,|\, \omega,\, \theta = i)\, \sigma(\omega)\, dF(\omega)} =$$

$$= \rho \frac{\int \mathbb{P}(s = b \,|\, \omega,\, \theta = h)\, \sigma(\omega)\, dF(\omega)}{\sum_{i \in \{l,h\}} \mathbb{P}(\theta = i) \int \mathbb{P}(s = b \,|\, \omega,\, \theta = i)\, \sigma(\omega)\, dF(\omega)}$$

$$\gamma^l \equiv \mathbb{P}(\theta = h \,|\, x = b,\, s = g) = \frac{\mathbb{P}(\theta = h,\, x = b,\, s = g)}{\mathbb{P}(x = b,\, s = g)} =$$

$$= \frac{\int \overbrace{\mathbb{P}(\theta = h \,|\, \omega)}^{=\mathbb{P}(\theta=h)=\rho} \mathbb{P}(x = b,\, s = g \,|\, \omega,\, \theta = h)\, \sigma(\omega)\, dF(\omega)}{\sum_{i \in \{l,\, h\}} \int \underbrace{\mathbb{P}(\theta = i \,|\, \omega)}_{=\mathbb{P}(\theta=i)} \mathbb{P}(x = b,\, s = g \,|\, \omega,\, \theta = i)\, \sigma(\omega)\, dF(\omega)} =$$

$$= \rho \, \frac{\int \mathbb{P}(x = b,\, s = g \,|\, \omega,\, \theta = h)\, \sigma(\omega)\, dF(\omega)}{\sum_{i \in \{l,\, h\}} \mathbb{P}(\theta = i) \int \mathbb{P}(x = b,\, s = g \,|\, \omega,\, \theta = i)\, \sigma(\omega)\, dF(\omega)} =$$

$$= \rho \, \frac{\int \mathbb{P}(s = g \,|\, x = b,\, \omega,\, \theta = h)\, \overbrace{\mathbb{P}(x = b \,|\, \omega,\, \theta = h)}^{\mathbb{P}(x=b \,|\, \omega)=1-p(\omega)}\, \sigma(\omega)\, dF(\omega)}{\sum_{i \in \{l,\, h\}} \mathbb{P}(\theta = i) \int \mathbb{P}(s = g \,|\, x = b,\, \omega,\, \theta = i)\, \underbrace{\mathbb{P}(x = b \,|\, \omega,\, \theta = i)}_{\mathbb{P}(x=b \,|\, \omega)=1-p(\omega)}\, \sigma(\omega)\, dF(\omega)}$$

Proof of Proposition 1

The expected utility of the agent with signal ω when choosing the *risky* project is

$$(1-\lambda)\kappa \underbrace{(p(\omega)\alpha(v_r - c) - (1-p(\omega))(1-\beta)c)}_{\pi(risk \mid \omega)} + \lambda \mathbb{E}(\gamma \mid \omega)$$

We define ω^* as the signal that equalizes the managerial expected utilities when choosing the *safe* or the *risky* project

$$(1-\lambda)\kappa v_s + \lambda\rho = (1-\lambda)\kappa \underbrace{(p(\omega^*)\alpha(v_r - c) - (1-p(\omega^*))(1-\beta)c)}_{\pi(risk \mid \omega^*)} + \lambda \mathbb{E}(\gamma \mid \omega^*)$$

It is sufficient to prove that the expected utility when choosing the *risky* project is an increasing function of ω. Notice that as $p(\omega)$ is increasing, the expected payoff $\pi(risk \mid \omega)$ is also growing with ω. For the whole utility to be increasing in ω, either $\mathbb{E}(\gamma \mid \omega)$ must be an increasing function of ω or the positive effect on $\pi(risk \mid \omega)$ must be larger than the negative one on $\mathbb{E}(\gamma \mid \omega)$. We study separately the sufficient conditions for these two cases.

Case 1: We can rewrite the expected reputation of a generic agent whose signal is ω, when choosing the risky project, as

$$\mathbb{E}(\gamma \mid \omega) \equiv p(\omega)\rho\alpha_h + (1-p(\omega))\rho(1-\beta_h) +$$

$$+(p(\omega)(1-\alpha) + (1-p(\omega))\beta) \underbrace{\frac{\int_{\sigma(l)=1} \rho((1-\alpha_h)p(l) + \beta_h(1-p(l)))dF(l)}{\int_{\sigma(l)=1}(1-\alpha)p(l) + \beta(1-p(l))dF(l)}}_{\equiv \gamma^0(risky)}$$

We can thus rewrite the expected reputation as the sum of two components: one independent on ω, the other one dependent on it

$$\overbrace{\rho(1-\beta_h)+\beta\gamma^0(risky)}^{\text{independent of }\omega}+p(\omega)(\rho\underbrace{(\alpha_h+\beta_h-1)}_{>0}+\gamma^0(risky)\underbrace{(1-\alpha-\beta)}_{<0})$$

As $p(\omega)$, is increasing, the expected reputation is increasing in ω if and only if

$$\rho(\alpha_h+\beta_h-1)+\gamma^0(w^*)(1-\alpha-\beta)>0$$

Using the definition of $\gamma^0(w^*)$, we can express this condition as:

$$\rho(\alpha_h+\beta_h-1)+\frac{\int_{\sigma(l)=1}\rho((1-\alpha_h)p(l)+\beta_h(1-p(l))dF(l)}{\int_{\sigma(l)=1}(1-\alpha)p(l)+\beta(1-p(l))dF(l)}(1-\alpha-\beta)>0$$

This is equivalent to

$$\underbrace{\frac{\rho}{\int_{\sigma(l)=1}(1-\alpha)p(l)+\beta(1-p(l))dF(l)}}_{>0}\times$$

$$\times\left(\int_{\sigma(l)=1}(\alpha_h+\beta_h-1)((1-\alpha)p(l)+\beta(1-p(l)))dF(l)+\right.$$

$$\left.+\int_{\sigma(l)=1}(1-\alpha-\beta)((1-\alpha_h)p(l)+\beta_h(1-p(l)))dF(l)\right)>0$$

This is true if and only if

$$\frac{1-\alpha_h}{\beta_h}<\frac{1-\alpha}{\beta}.$$

Case 2: When the expected reputation is decreasing in ω, we can nonetheless look

for conditions that guarantee that the positive derivative of the expected payoff with respect to ω, when choosing the *risky* project, dominates.
In a similar way to what we did before, we express the portion of the managerial utility depending on his expected reputation as

$$\lambda\Big(\overbrace{(1-\beta_h)\rho+\beta\gamma^0(\omega^*)}^{\text{independent of }\omega}+p(\omega)(\rho\overbrace{(\alpha_h+\beta_h-1)}^{>0}+\gamma^0(\omega^*)\overbrace{(1-\alpha-\beta)}^{<0})\Big)$$

When $\rho(\alpha_h+\beta_h-1)+\gamma^0(w^*)(1-\alpha-\beta)<0$, the derivative of this expression with respect to ω cannot be lower than $\lambda p'(\omega)(\rho(\alpha_h+\beta_h-1)+(1-\alpha-\beta))$, as $\gamma^0(w^*)\in[\,0,1]$.
The derivative of the part of the utility function related to the return on the risky project is, instead,

$$(1-\lambda)\kappa(p'(\omega)\alpha(v_r-c)+p'(\omega)(1-\beta)c)>0.$$

For the whole managerial utility to be an increasing function of ω it is then sufficient that

$$(1-\lambda)\kappa(p'(\omega)\alpha(v_r-c)+p'(\omega)(1-\beta)c>-\lambda p'(\omega)\quad\underbrace{(\rho(\alpha_h+\beta_h-1)+(1-\alpha-\beta))}_{\equiv\rho\alpha_h+\rho\beta_h-\rho+1-\rho\alpha_h-(1-\rho)\alpha_l-\rho\beta_h-(1-\rho)\beta_l}$$

that is, whenever:

$$\frac{\lambda}{1-\lambda}<\frac{\kappa(\alpha(v_r-c)+(1-\beta)c)}{(1-\rho)(\alpha_l+\beta_l-1)}.$$

To sum up, whenever $\frac{1-\alpha_h}{\beta_h}<\frac{1-\alpha}{\beta}$ the manager is better off choosing the *risky* project if and only if his signal ω is bigger than the equilibrium ω^*. When $\frac{1-\alpha_h}{\beta_h}>\frac{1-\alpha}{\beta}$, this is also true if career concerns are not too strong, that is λ is small enough.

■

Proof of Lemma 1

Consider two cutoff equilibria characterized by thresholds w_1^* and w_2^*, with $w_1^* > w_2^*$. Let us consider the conditions that guarantee that $\gamma^0(w_1^*) > \gamma^0(w_2^*)$. These are the beliefs in case a *risky* project is abandoned, under the two equilibria. We study in which circumstances the following holds:

$$\gamma^0(w_1^*) \equiv \frac{\int_{w_1^*}^{\overline{w}} \rho((1-\alpha_h)p(w) + \beta_h(1-p(w)))dF(w)}{\int_{w_1^*}^{\overline{w}} ((1-\alpha)p(w) + \beta(1-p(w)))dF(w)}$$

$$>$$

$$\frac{\int_{w_2^*}^{\overline{w}} \rho((1-\alpha_h)p(w) + \beta_h(1-p(w)))dF(w)}{\int_{w_2^*}^{\overline{w}} ((1-\alpha)p(w) + \beta(1-p(w)))dF(w)} \equiv \gamma^0(w_2^*)$$

As the two denominators are non negative, this is equivalent to:

$$(\int_{w_2^*}^{\overline{w}} ((1-\alpha)p(w) + \beta(1-p(w)))dF(w))(\int_{w_1^*}^{\overline{w}} ((1-\alpha_h)p(w) + \beta_h(1-p(w)))dF(w))$$

$$>$$

$$(\int_{w_1^*}^{\overline{w}} ((1-\alpha)p(w) + \beta(1-p(w)))dF(w))(\int_{w_2^*}^{\overline{w}} ((1-\alpha_h)p(w) + \beta_h(1-p(w)))dF(w))$$

Using the definitions of α and β, that is, $\alpha \equiv \rho\alpha_h + (1-\rho)\alpha_l$ and $\beta \equiv \rho\beta_h + (1-\rho)\beta_l$, we can rewrite this expression as:

$$((1-\alpha)\beta_h - \beta(1-\alpha_h)) \int_{w_2^*}^{\overline{w}} p(w)dF(w) \int_{w_1^*}^{\overline{w}} (1-p(w))dF(w)$$

$$>$$

$$((1-\alpha)\beta_h - \beta(1-\alpha_h)) \int_{w_1^*}^{\overline{w}} p(w)dF(w) \int_{w_2^*}^{\overline{w}} (1-p(w))dF(w)$$

Suppose now that $(1-\alpha)\beta_h \leq \beta(1-\alpha_h)$ - which is equivalent to $(1-\alpha_l)\beta_h \leq \beta_l(1-\alpha_h)$). Then the inequality holds if and only if:

$$\int_{w_2^*}^{\overline{w}} p(w)dF(w) \int_{w_1^*}^{\overline{w}} (1-p(w))dF(w) < \int_{w_1^*}^{\overline{w}} p(w)dF(w) \int_{w_2^*}^{\overline{w}} (1-p(w))dF(w)$$

that is, if and only if

$$\int_{w_2^*}^{\overline{w}} p(w)dF(w) \int_{w_1^*}^{\overline{w}} dF(w) - \int_{w_2^*}^{\overline{w}} p(w)dF(w) \int_{w_1^*}^{\overline{w}} p(w)dF(w)$$

$$<$$

$$\int_{w_1^*}^{\overline{w}} p(w)dF(w) \int_{w_2^*}^{\overline{w}} dF(w) - \int_{w_1^*}^{\overline{w}} p(w)dF(w) \int_{w_2^*}^{\overline{w}} p(w)dF(w)$$

As the second terms on each side of the inequality are the same, this simplifies to:

$$(1 - F(w_1^*)) \int_{w_2^*}^{\overline{w}} p(w)dF(w) < (1 - F(w_2^*)) \int_{w_1^*}^{\overline{w}} p(w)dF(w).$$

Because $p(w) < p(w_1^*)$ for any $w < w_1^*$, notice that the left hand side of this inequality is at most:

$$(1 - F(w_1^*))(F(w_1^*) - F(w_2^*))p(w_1^*) + (1 - F(w_1^*)) \int_{w_1^*}^{\overline{w}} p(w)dF(w) - \epsilon_1$$

for some $\epsilon_1 > 0$. Therefore, the inequality necessarily holds if the following holds:

$$(1 - F(w_1^*))(F(w_1^*) - F(w_2^*))p(w_1^*) - \epsilon_1 < (F(w_1^*) - F(w_2^*)) \int_{w_1^*}^{\overline{w}} p(w)dF(w).$$

The right hand side of this equation cannot be lower than $(F(w_1^*) - F(w_2^*))(1 - F(w_1^*))p(w_1^*) + \epsilon_2$ for some $\epsilon_2 > 0$. Therefore this condition always holds and $\gamma^0(w_1^*) > \gamma^0(w_2^*)$ when $(1 - \alpha_l)\beta_h \leq \beta_l(1 - \alpha_h)$.

With a similar argument we could show that the opposite is true when $(1 - \alpha_l)\beta_h \geq \beta_l(1 - \alpha_h)$. ■

Proof of Corollary 1

Let ω_1^* and ω_2^* be two equilibrium cutoffs, such that $\omega_1^* > \omega_2^*$. By definition it must be the case that the expected utilities of the two *marginal* individuals choosing the *risky* project equals the utility when choosing the *safe* one

$$(1-\lambda)\kappa(v_s - c) + \lambda\rho$$

$$=$$

$$(1-\lambda)\kappa((p(\omega_1^*)\alpha(v_r - c) + (1-p(\omega_1))(1-\beta)(-c) + \lambda(p(\omega_1^*)\rho\alpha_h + (1-p(\omega_1^*))\rho(1-\beta_h)+$$

$$+(p(\omega_1^*)(1-\alpha) + (1-p(\omega_1^*))\beta)\frac{\int_{w_1^*}^{\overline{w}} \rho((1-\alpha_h)p(\omega) + \beta_h(1-p(\omega)))dF(\omega)}{\int_{w_1^*}^{\overline{w}}((1-\alpha)p(\omega) + \beta(1-p(\omega)))dF(\omega)})$$

$$=$$

$$(1-\lambda)\kappa((p(\omega_2^*)\alpha(v_r - c) + (1-p(\omega_2^*))(1-\beta)(-c) + \lambda(p(\omega_2^*)\rho\alpha_h + (1-p(\omega_2^*))\rho(1-\beta_h)+$$

$$+(p(\omega_2^*)(1-\alpha) + (1-p(\omega_2^*))\beta)\frac{\int_{w_2^*}^{\overline{w}} \rho((1-\alpha_h)p(\omega) + \beta_h(1-p(\omega)))dF(\omega)}{\int_{w_2^*}^{\overline{w}}((1-\alpha)p(\omega) + \beta(1-p(\omega)))dF(\omega)})$$

For this to hold, it must be the case that:

$$\frac{(1-\lambda)}{\lambda}\kappa((p(\omega_2^*)-p(\omega_1^*))\alpha(v_r-c)+(p(\omega_2^*)-p(\omega_1^*))(1-\beta)c)+$$

$$+(p(\omega_2^*)-p(\omega_1^*))\rho\alpha_h-(p(\omega_2^*)-p(\omega_1^*))\rho(1-\beta_h)$$

$$=$$

$$(p(\omega_1^*)(1-\alpha)+(1-p(\omega_1^*))\beta)\frac{\int_{w_1^*}^{\overline{w}}\rho((1-\alpha_h)p(\omega)+\beta_h(1-p(\omega)))dF(w)}{\int_{w_1^*}^{\overline{w}}((1-\alpha)p(\omega)+\beta(1-p(\omega)))dF(\omega)}+$$

$$-(p(\omega_2^*)(1-\alpha)+(1-p(\omega_2^*))\beta)\frac{\int_{w_2^*}^{\overline{w}}\rho((1-\alpha_h)p(\omega)+\beta_h(1-p(\omega)))dF(\omega)}{\int_{w_2^*}^{\overline{w}}((1-\alpha)p(\omega)+\beta(1-p(\omega)))dF(\omega)} \qquad \text{(C.1)}$$

Notice that the left hand side of equation (C.1) is negative, as $p(\omega_2^*) < p(\omega_1^*)$ and $\alpha_h+\beta_h-1>0$.

As $p(\omega_1^*) \equiv p(\omega_2^*)+(p(\omega_1^*)-p(\omega_2^*)) > p(\omega_2^*)$, we can now rewrite the right hand side of (C.1) as:

$$(p(\omega_2^*)(1-\alpha)+(1-p(\omega_2^*))\beta+(p(\omega_1^*)-p(\omega_2^*))(1-\alpha-\beta)))\times$$

$$\times\frac{\int_{w_1^*}^{\overline{w}}\rho((1-\alpha_h)p(\omega)+\beta_h(1-p(\omega)))dF(\omega)}{\int_{w_1^*}^{\overline{w}}((1-\alpha)p(\omega)+\beta(1-p(\omega)))dF(\omega)}+$$

$$-(p(\omega_2^*)(1-\alpha)+(1-p(\omega_2^*))\beta)\frac{\int_{w_2^*}^{\overline{w}}\rho((1-\alpha_h)p(\omega)+\beta_h(1-p(\omega)))dF(\omega)}{\int_{w_2^*}^{\overline{w}}((1-\alpha)p(\omega)+\beta(1-p(\omega)))dF(\omega)}$$

and then as:

$$(p(\omega_1^*)-p(\omega_2^*))(1-\alpha-\beta))\frac{\int_{w_1^*}^{\overline{w}}\rho((1-\alpha_h)p(\omega)+\beta_h(1-p(\omega)))dF(\omega)}{\int_{w_1^*}^{\overline{w}}((1-\alpha)p(\omega)+\beta(1-p(\omega)))dF(\omega)}+$$

$$+(p(\omega_2^*)(1-\alpha)+(1-p(\omega_2^*))\beta)\times$$

$$\times(\frac{\int_{w_1^*}^{\overline{w}}\rho((1-\alpha_h)p(\omega)+\beta_h(1-p(\omega)))dF(\omega)}{\int_{w_1^*}^{\overline{w}}((1-\alpha)p(\omega)+\beta(1-p(\omega)))dF(\omega)}-\frac{\int_{w_2^*}^{\overline{w}}\rho((1-\alpha_h)p(\omega)+\beta_h(1-p(\omega)))dF(\omega)}{\int_{w_2^*}^{\overline{w}}((1-\alpha)p(\omega)+\beta(1-p(\omega)))dF(\omega)})$$

The first part of this term is negative as $\alpha+\beta>1$ and it is bounded below by $(p(\omega_1^*)-p(\omega_2^*))(1-\alpha-\beta)$, as $\gamma_0(\omega_1^*)\in[0,1]$.

By the previous Lemma the second part is positive if and only if $(1-\alpha_l)\beta_h\leq\beta_l(1-\alpha_h)$. In this scenario, the right hand side of equation (4) has a negative component and a positive one. Thus, if $(1-\alpha_l)\beta_h\leq\beta_l(1-\alpha_h)$, (4) cannot hold whenever:

$$(1-\lambda)\kappa((p(\omega_2^*)-p(\omega_1^*))\alpha(v_r-c)+(p(\omega_2^*)-p(\omega_1^*)(1-\beta)c)+$$

$$+\lambda((p(\omega_2^*)-p(\omega_1^*))\rho\alpha_h-(p(\omega_2^*)-p(\omega_1^*))\rho(1-\beta_h))$$

$$<$$

$$\lambda(p(\omega_2^*) - p(\omega_1^*))(\alpha + \beta - 1).$$

This is equivalent to:

$$(1-\lambda)\kappa(\alpha(v_r - c) + (1-\beta)c) + \lambda\rho(\alpha_h + \beta_h - 1)) > \lambda(\alpha + \beta - 1)$$

and, after simplification, to:

$$\frac{\lambda}{1-\lambda} < \frac{\kappa(\alpha(v_r - c) + (1-\beta)c)}{(1-\rho)(\alpha_l + \beta_l - 1)}.$$

■

Additional beliefs with N agents

$$\gamma_N^{nl} \equiv \mathbb{P}(\theta_N = h \,|\, x = b,\; s = b) = \frac{\int_{\omega^*}^{\overline{\omega}} \mathbb{P}(\theta_N = h,\; x = b,\; s = b \,|\omega)\sigma(\omega)dF(\omega)}{\int_{\omega^*}^{\overline{\omega}} \mathbb{P}(x = b,\; s = b \,|\omega)\sigma(\omega)dF(\omega)} =$$

$$= \frac{\int_{\omega^*}^{\overline{\omega}} \overbrace{\mathbb{P}(\theta_N = h \,|\, \omega)}^{=\mathbb{P}(\theta=h)=\rho} \, \mathbb{P}(x = b,\; s = b \,|\, \omega,\; \theta_N = h)\, \sigma(\omega)\, dF(\omega)}{\sum_{i \in \{l,h\}} \int_{\omega^*}^{\overline{\omega}} \underbrace{\mathbb{P}(\theta_N = i \,|\, \omega)}_{=\mathbb{P}(\theta_N = i)} \mathbb{P}(x = b,\; s = b \,|\, \omega,\; \theta_N = i)\, \sigma(\omega)\, dF(\omega)} =$$

$$= \rho \, \frac{\int_{\omega^*}^{\overline{\omega}} \mathbb{P}(x = b,\; s = b \,|\, \omega,\; \theta_N = h)\, \sigma(\omega)\, dF(\omega)}{\sum_{i \in \{l,h\}} \mathbb{P}(\theta_N = i) \int_{\omega^*}^{\overline{\omega}} \mathbb{P}(x = b,\; s = b \,|\, \omega,\; \theta_N = i)\, \sigma(\omega)\, dF(\omega)} =$$

$$= \rho \, \frac{\int_{\omega^*}^{\overline{\omega}} \mathbb{P}(s = b \,|\, x = b,\; \omega,\; \theta = h)\; \overbrace{\mathbb{P}(x = b \,|\, \omega,\; \theta_N = h)}^{=\mathbb{P}(x=b \,|\, \omega)=p(\omega)}\; \sigma(\omega)\, dF(\omega)}{\sum_{i \in \{l,h\}} \mathbb{P}(\theta_N = i) \int_{\omega^*}^{\overline{\omega}} (s = b \,|\, x = b,\; \omega,\; \theta_N = i)\; \underbrace{\mathbb{P}(x = b \,|\, \omega,\; \theta_N = i)}_{=\mathbb{P}(x=b \,|\, \omega)=p(\omega)}\; \sigma(\omega)\, dF(\omega)} = \rho \frac{\beta_h}{\beta}$$

$$\gamma_N^0 \equiv \mathbb{P}(\theta_N = h \,|\, s_N = b,\; s_1 = b, ..., s_{N-1} = b) = \frac{\mathbb{P}(\theta_N = h,\; s_1 = b, ..., s_N = b)}{\mathbb{P}(s_1 = b, ..., s_N = b)} =$$

$$= \frac{\int_{\omega^*}^{\overline{\omega}} \mathbb{P}(\theta_N = h, s_1 = b, ..., s_N = b \,|\omega)\sigma(\omega)dF(\omega)}{\int_{\omega^*}^{\overline{\omega}} \mathbb{P}(s_1 = b, ..., s_N = b \,|\omega)\sigma(\omega)dF(\omega)} =$$

$$= \frac{\int_{\omega^*}^{\varpi} \overbrace{\mathbb{P}(\theta_N = h \,|\, \omega)}^{=\mathbb{P}(\theta_N = h) = \rho} \, \mathbb{P}(s_1 = b, ..., s_N = b \,|\, \omega,\, \theta_N = h)\, \sigma(\omega)\, dF(\omega)}{\sum_{i \in \{l,h\}} \int_{\omega^*}^{\varpi} \underbrace{\mathbb{P}(\theta_N = i \,|\, \omega)}_{=\mathbb{P}(\theta_N = i)} \mathbb{P}(s_1 = b, ..., s_N = b \,|\, \omega,\, \theta_N = i)\, \sigma(\omega)\, dF(\omega)} =$$

Notice that $\mathbb{P}(s_1 = b, .., s_N = b \,|\, \omega, \theta = i)$ can be computed as:

$$\begin{aligned} &\mathbb{P}(s_1 = b, .., s_N = b \,|\, x = g, \omega,\, \theta = i)\mathbb{P}(x = g|\omega, \theta_N = i) \quad + \\ &+\mathbb{P}(s_1 = b, .., s_N = b|x = b, \omega, \theta = i)\mathbb{P}(x = b|\omega, \theta_N = i) \quad . \end{aligned} \tag{C.2}$$

Now, the first term of equation (5) - $\mathbb{P}(s_1 = b, ..., s_N = b \,|\, x = g,\, \omega,\, \theta_N = i)$ - is:

$$\mathbb{P}(s_N = b \,|s_1 = b, ..., s_{N-1} = b,\, x = g,\, \omega,\, \theta_N = i)\mathbb{P}(s_1 = b, ..., s_{N-1} = b \,|\, x = g,\, \omega,\, \theta_N = i)\times$$

$$\times\mathbb{P}(s_N = b \,|\, x = g,\, \theta_N = i)\mathbb{P}(s_{N-1} = b \,|\, x = g)...\mathbb{P}(s_1 = b \,|\, x = g) =$$

$$= (1 - \alpha_i)(1 - \alpha)^{N-1}.$$

Similarly, the second term of (5) can be computed as $\mathbb{P}(s_1 = b, ..., s_N = b \,|\, x = b,\, \omega,\, \theta_N = i) = \beta_i \beta^{N-1}$. Thus, we have:

$$\gamma_N^0 = \rho \frac{(1 - \alpha_h)(1 - \alpha)^{N-1}\tilde{p} + \beta_h \beta^{N-1}(1 - \tilde{p})}{(1 - \alpha)^N \tilde{p} + \beta^N (1 - \tilde{p})}$$

$$\gamma_N^{nr} \equiv \mathbb{P}(\theta_N = h \,|\, x = g,\, s = b) = \frac{\int_{\omega^*}^{\varpi} \mathbb{P}(\theta_N = h,\, x = g,\, s = b \,|\omega)\sigma(\omega) dF(\omega)}{\int_{\omega^*}^{\varpi} \mathbb{P}(x = g,\, s = b \,|\omega)\sigma(\omega) dF(\omega)} =$$

$$= \frac{\int_{\omega^*}^{\overline{\omega}} \overbrace{\mathbb{P}(\theta_N = h \,|\, \omega)}^{=\mathbb{P}(\theta_N = h) = \rho} \, \mathbb{P}(x = g,\, s = b \,|\, \omega,\, \theta_N = h)\, \sigma(\omega)\, dF(\omega)}{\sum_{i \in \{l,\, h\}} \int_{\omega^*}^{\overline{\omega}} \underbrace{\mathbb{P}(\theta_N = i \,|\, \omega)}_{=\mathbb{P}(\theta_N = i)} \mathbb{P}(x = g,\, s = b \,|\, \omega,\, \theta = i)\, \sigma(\omega)\, dF(\omega)} =$$

$$= \rho \, \frac{\int_{\omega^*}^{\overline{\omega}} \mathbb{P}(x = g,\, s = b \,|\, \omega,\, \theta_N = h)\, \sigma(\omega)\, dF(\omega)}{\sum_{i \in \{l,\, h\}} \mathbb{P}(\theta_N = i) \int_{\omega^*}^{\overline{\omega}} \mathbb{P}(x = g,\, s = b \,|\, \omega,\, \theta_N = i)\, \sigma(\omega)\, dF(\omega)} =$$

$$= \rho \, \frac{\int_{\omega^*}^{\overline{\omega}} \mathbb{P}(s = b \,|\, x = g,\, \omega,\, \theta = h) \, \overbrace{\mathbb{P}(x = g \,|\, \omega,\, \theta_N = h)}^{\mathbb{P}(x = g \,|\, \omega) = p(\omega)} \, \sigma(\omega)\, dF(\omega)}{\sum_{i \in \{l,\, h\}} \mathbb{P}(\theta_N = i) \, \int_{\omega^*}^{\overline{\omega}} \mathbb{P}(s = b \,|\, x = g,\, \omega,\, \theta_N = i) \, \underbrace{\mathbb{P}(x = g \,|\, \omega,\, \theta_N = i)}_{\mathbb{P}(x = g \,|\, \omega) = p(\omega)} \, \sigma(\omega)\, dF(\omega)} =$$

$$= \rho \frac{1 - \alpha_h}{1 - \alpha}.$$

Proof of Proposition 3

We start showing that Lemma 1 and Lemma 2 also hold in the generalized version of the model.

Consider, again, two cutoff equilibria characterized by thresholds ω_1^* and ω_2^*, with $\omega_1^* > \omega_2^*$. We want to show that $\gamma^0(\omega_1^*) > \gamma^0(\omega_2^*)$ if and only if $\frac{1-\alpha^H}{\beta^H} > \frac{1-\alpha}{\beta}$. In doing this, we use the new definition of belief in case of termination of the *risky* project. $\gamma^0(\omega_1^*) > \gamma^0(\omega_2^*)$ when:

$$\rho\frac{\int_{\omega_1^*}^{\overline{\omega}}(1-\alpha_h)(1-\alpha)^{N-1}p(\omega)+\beta_h\beta^{N-1}(1-p(\omega))dF(\omega)}{\int_{\omega_1^*}^{\overline{\omega}}((1-\alpha)^N p(\omega)+\beta^N(1-p(\omega)))dF(\omega)}$$

$$>$$

$$\rho\frac{\int_{\omega_2^*}^{\overline{\omega}}(1-\alpha_h)(1-\alpha)^{N-1}p(\omega)+\beta_h\beta^{N-1}(1-p(\omega))dF(\omega)}{\int_{\omega_2^*}^{\overline{\omega}}(1-\alpha)^N p(\omega)+\beta^N(1-p(\omega))dF(\omega)}$$

This is equivalent to:

$$(1-\alpha_h)(1-\alpha)^{N-1}\beta^N\int_{\omega_1^*}^{\overline{\omega}}p(\omega)dF(\omega)\int_{\omega_2^*}^{\overline{\omega}}(1-p(\omega))dF(\omega)+$$
$$+(1-\alpha)^N\beta_h\beta^{N-1}\int_{\omega_1^*}^{\overline{\omega}}(1-p(\omega))dF(\omega)\int_{\omega_2^*}^{\overline{\omega}}p(\omega)dF(\omega)$$

$$>$$

$$(1-\alpha_h)(1-\alpha)^{N-1}\beta^N \int_{\omega_1^*}^{\overline{\omega}} p(\omega)dF(\omega) \int_{\omega_2^*}^{\overline{\omega}} (1-p(w))dF(\omega) + \\ + (1-\alpha)^N \beta_h \beta^{N-1} \int_{\omega_1^*}^{\overline{\omega}} (1-p(\omega))dF(\omega) \int_{\omega_2^*}^{\overline{\omega}} p(\omega)dF(\omega)$$

We now divide everything by $(1-\alpha)^{N-1}\beta^{N-1}$ and rearrange, to get:

$$((1-\alpha)\beta_h - \beta(1-\alpha_h)) \int_{\omega_2^*}^{\overline{\omega}} p(\omega)dF(\omega) \int_{w_1^*}^{\overline{\omega}} (1-p(\omega))dF(\omega)$$

$$>$$

$$((1-\alpha)\beta_h - \beta(1-\alpha_h)) \int_{w_1^*}^{\overline{\omega}} p(\omega)dF(\omega) \int_{\omega_2^*}^{\overline{\omega}} (1-p(\omega))dF(\omega)$$

This is equivalent to what we had in the proof of Lemma 1. Therefore for ω_1^* and ω_2^*, with $w_1^* > w_2^*$, we have that $\gamma^0(w_1^*) > \gamma^0(w_2^*)$ if and only if $\frac{1-\alpha_h}{\beta_h} > \frac{1-\alpha}{\beta}$.

Now we need to show that, as in Lemma 2, upon observing ω^* the marginal agent is less optimistic than the principal in evaluating his own ability if and only if $\frac{1-\alpha_h}{\beta_h} > \frac{1-\alpha}{\beta}$. The proof of this claim is straightforward. Suppose $\frac{1-\alpha_h}{\beta_h} > \frac{1-\alpha}{\beta}$.

Notice that, for the marginal agent:

$$E(\gamma \mid \omega^*) \equiv p(\omega^*)\alpha\rho\frac{\alpha_h}{\alpha} + ((1 - p(\omega^*)(1-\beta))\rho\frac{(1-\beta_h)}{(1-\beta)} +$$

$$+p(\omega^*)(1-\alpha-(1-\alpha)^N)\rho\frac{1-\alpha_h}{1-\alpha} + (1-p(\omega^*))(1-(1-\beta)-\beta^N)\rho\frac{\beta_h}{\beta} +$$

$$+(p(\omega^*)(1-\alpha)^N + (1-p(\omega^*))\beta^N)\rho\,\frac{(1-\alpha_h)(1-\alpha)^{N-1}\tilde{p} + \beta_h\beta^{N-1}(1-\tilde{p})}{(1-\alpha)^N\tilde{p} + \beta^N(1-\tilde{p})}$$

$$>$$

$$p(\omega^*)\alpha\rho\frac{\alpha_h}{\alpha} + ((1 - p(\omega^*)(1-\beta))\rho\frac{(1-\beta_h)}{(1-\beta)} +$$

$$+p(\omega^*)(1-\alpha-(1-\alpha)^N)\rho\frac{1-\alpha_h}{1-\alpha} + (1-p(\omega^*))(1-(1-\beta)-\beta^N)\rho\frac{\beta_h}{\beta} +$$

$$+(p(\omega^*)(1-\alpha)^N + (1-p(\omega^*))\beta^N)\rho\,\frac{(1-\alpha_h)(1-\alpha)^{N-1}p(\omega^*) + \beta_h\beta^{N-1}(1-p(\omega^*))}{(1-\alpha)^N p(\omega^*) + \beta^N(1-p(\omega^*))} = \rho$$

Hence, the marginal agent takes too much risk if and only if $\frac{1-\alpha_h}{\beta_h} > \frac{1-\alpha}{\beta}$. ■

Proof of Proposition 4

We start from the definition of $E(\gamma_N \mid \omega^*)$ and $E(\gamma_{N+1} \mid \omega^*)$. These are, respectively:

$$
\begin{aligned}
E(\gamma_N \mid \omega^*) \;\equiv\;& p(\omega^*)\alpha\rho\frac{\alpha_h}{\alpha} + ((1-p(\omega^*)(1-\beta))\rho\frac{(1-\beta_h)}{(1-\beta)} + \\
&+p(\omega^*)(1-\alpha-(1-\alpha)^N)\rho\frac{1-\alpha_h}{1-\alpha} + (1-p(\omega^*))(1-(1-\beta)-\beta^N)\rho\frac{\beta_h}{\beta} + \\
&+(p(\omega^*)(1-\alpha)^N + (1-p(\omega^*))\beta^N)\rho\,\frac{(1-\alpha_h)(1-\alpha)^{N-1}\tilde{p} + \beta_h\beta^{N-1}(1-\tilde{p})}{(1-\alpha)^N\tilde{p} + \beta^N(1-\tilde{p})}
\end{aligned}
$$

and:

$$
\begin{aligned}
E(\gamma_{N+1} \mid \omega^*) \;\equiv\;& p(\omega^*)\alpha\rho\frac{\alpha_h}{\alpha} + ((1-p(\omega^*)(1-\beta))\rho\frac{(1-\beta_h)}{(1-\beta)} + \\
&+p(\omega^*)(1-\alpha-(1-\alpha)^{N+1})\rho\frac{1-\alpha_h}{1-\alpha} + (1-p(\omega^*))(1-(1-\beta)-\beta^{N+1})\rho\frac{\beta_h}{\beta} + \\
&+(p(\omega^*)(1-\alpha)^{N+1} + (1-p(\omega^*))\beta^{N+1})\rho\,\frac{(1-\alpha_h)(1-\alpha)^{N}\tilde{p} + \beta_h\beta^{N}(1-\tilde{p})}{(1-\alpha)^{N+1}\tilde{p} + \beta^{N+1}(1-\tilde{p})}
\end{aligned}
$$

We want to show that $E(\gamma_N \mid \omega^*) > E(\gamma_{N+1} \mid \omega^*)$ if and only if $\frac{1-\alpha_h}{\beta_h} > \frac{1-\alpha_l}{\beta_l}$. Noticing that the first two addends in the definitions of $E(\gamma_N \mid \omega^*)$ and $E(\gamma_{N+1} \mid \omega^*)$ coincide and dividing everything by ρ, $E(\gamma_N \mid \omega^*) > E(\gamma_{N+1} \mid \omega^*)$ if and only if

$$p(\omega^*)(1-\alpha-(1-\alpha)^N)\frac{1-\alpha_h}{1-\alpha}+(1-p(\omega^*))(1-(1-\beta)-\beta^N)\frac{\beta_h}{\beta}+$$

$$+(p(\omega^*)(1-\alpha)^N+(1-p(\omega^*))\beta^N)\,\frac{(1-\alpha_h)(1-\alpha)^{N-1}\tilde{p}+\beta_h\beta^{N-1}(1-\tilde{p})}{(1-\alpha)^N\tilde{p}+\beta^N(1-\tilde{p})}$$

$$>$$

$$p(\omega^*)(1-\alpha-(1-\alpha)^{N+1})\frac{1-\alpha_h}{1-\alpha}+(1-p(\omega^*))(1-(1-\beta)-\beta^{N+1})\frac{\beta_h}{\beta}+$$

$$+(p(\omega^*)(1-\alpha)^{N+1}+(1-p(\omega^*))\beta^{N+1})\,\frac{(1-\alpha_h)(1-\alpha)^{N}\tilde{p}+\beta_h\beta^{N}(1-\tilde{p})}{(1-\alpha)^{N+1}\tilde{p}+\beta^{N+1}(1-\tilde{p})}$$

Now we multiply everything by $\left((1-\alpha)^N\tilde{p}+\beta^N(1-\tilde{p})\right)\left((1-\alpha)^{N+1}\tilde{p}+\beta^{N+1}(1-\tilde{p})\right)$ and adjust terms to get

$$(1-\alpha_h)\Bigg\{-p(\omega^*)(1-\alpha)^{N-1}\alpha\left((1-\alpha)^N\tilde{p}+\beta^N(1-\tilde{p})\right)\left((1-\alpha)^{N+1}\tilde{p}+\beta^{N+1}(1-\tilde{p})\right)+$$

$$+(1-\alpha)^{N-1}\tilde{p}\left((1-\alpha)^N p(\omega^*)+\beta^N(1-p(\omega^*))\right)\left((1-\alpha)^{N+1}\tilde{p}+\beta^{N+1}(1-\tilde{p})\right)+$$

$$-(1-\alpha)^N\tilde{p}\left((1-\alpha)^{N+1}p(\omega^*)+\beta^{N+1}(1-p(\omega^*))\right)\left((1-\alpha)^N\tilde{p}+\beta^N(1-\tilde{p})\right)\Bigg\}$$

$$>$$

$$\beta_h\Bigg\{(1-p(\omega^*))\beta^{N-1}(1-\beta)\left((1-\alpha)^N\tilde{p}+\beta^N(1-\tilde{p})\right)\left((1-\alpha)^{N+1}\tilde{p}+\beta^{N+1}(1-\tilde{p})\right)+$$

$$+\beta^N(1-\tilde{p})\left((1-\alpha)^{N+1}p(\omega^*)+\beta^{N+1}(1-p(\omega^*))\right)\left((1-\alpha)^N\tilde{p}+\beta^N(1-\tilde{p})\right)$$

$$-\beta^{N-1}(1-\tilde{p})\left((1-\alpha)^N p(\omega^*)+\beta^N(1-p(\omega^*))\right)\left((1-\alpha)^{N+1}\tilde{p}+\beta^{N+1}(1-\tilde{p})\right)\Bigg\}$$

We focus separately on the two sides of the inequality. From the left hand side we obtain:

$$(1-\alpha_h)\Bigg\{-p(\omega^*)(1-\alpha)^{N-1}\alpha\Big((1-\alpha)^{2N+1}\tilde{p}^2+(1-\alpha)^N\beta^{N+1}\tilde{p}(1-\tilde{p})+$$

$$+(1-\alpha)^{N+1}\beta^N\tilde{p}(1-\tilde{p})+\beta^{2N+1}(1-\tilde{p})^2\Big)+$$

$$+\tilde{p}(1-\alpha)^{N-1}\Big((1-\alpha)^{2N+1}p(\omega^*)\tilde{p}+(1-\alpha)^N\beta^{N+1}p(\omega^*)(1-\tilde{p})+$$

$$+(1-\alpha)^{N+1}\beta^N\tilde{p}(1-p(\omega^*))+\beta^{2N+1}(1-p(\omega^*))(1-\tilde{p})\Big)+$$

$$-\tilde{p}(1-\alpha)^N\Big((1-\alpha)^{2N+1}p(\omega^*)\tilde{p}+(1-\alpha)^{N+1}\beta^N p(\omega^*)(1-\tilde{p})+$$

$$+(1-\alpha)^N\beta^{N+1}\tilde{p}(1-p(\omega^*))+\beta^{2N+1}(1-p(\omega^*))(1-\tilde{p})\Big)\Bigg\}=$$

$$=(1-\alpha_h)\Bigg\{-\alpha(1-\alpha)^{2N-1}\beta^{N+1}p(\omega^*)\tilde{p}(1-\tilde{p})+$$

$$-\alpha(1-\alpha)^{2N}\beta^N p(\omega^*)\tilde{p}(1-\tilde{p})-\alpha(1-\alpha)^{N-1}\beta^{2N+1}p(\omega^*)(1-\tilde{p})^2+$$

$$(1-\alpha)^{2N-1}\beta^{N+1}\Big(p(\omega^*)\tilde{p}(1-\tilde{p})-(1-\alpha)(1-p(\omega^*))\tilde{p}^2\Big)+$$

$$+(1-\alpha)^{2N}\beta^N\Big(\tilde{p}^2(1-p(\omega^*))-(1-\alpha)p(\omega^*)\tilde{p}(1-\tilde{p})\Big)+$$

$$+(1-\alpha)^{N-1}\beta^{2N+1}\Big(\tilde{p}(1-p(\omega^*))(1-\tilde{p})-(1-\alpha)\tilde{p}(1-p(\omega^*))(1-\tilde{p})\Big)\Bigg\}=$$

$$=(1-\alpha_h)\Bigg\{(1-\alpha)^{2N-1}\beta^{N+1}\tilde{p}\Big(p(\omega^*)(1-\tilde{p})-(1-\alpha)(1-p(\omega^*))\tilde{p}-\alpha p(\omega^*)(1-\tilde{p})\Big)+$$

$$+(1-\alpha)^{2N}\beta^N\tilde{p}\Big(\tilde{p}(1-p(\omega^*)-(1-\alpha)p(\omega^*)(1-\tilde{p})-\alpha p(\omega^*)(1-\tilde{p})\Big)+$$

$$+(1-\alpha)^{N-1}\beta^{2N+1}\Big(\tilde{p}(1-p(\omega^*))-(1-\alpha)\tilde{p}(1-p(\omega^*))-\alpha p(\omega^*)(1-\tilde{p})\Big)\Bigg\}=$$

$$=(1-\alpha_h)\Bigg\{(1-\alpha)^{2N-1}\beta^{N+1}\tilde{p}\Big((1-\alpha)(p(\omega^*)-\tilde{p})\Big)+(1-\alpha)^{2N}\beta^N\tilde{p}\Big(\tilde{p}-p(\omega^*)\Big)+$$

$$+(1-\alpha)^{N-1}\beta^{2N+1}\Big(\alpha(\tilde{p}-p(\omega^*))\Big)\Bigg\}=$$

$$=(1-\alpha_h)\beta(\tilde{p}-p(\omega^*))\Bigg\{(1-\alpha)^{2N}\beta^{N-1}(1-\beta)\tilde{p}+(1-\alpha)^{N-1}\beta^N(1-\tilde{p})\alpha\Bigg\}.$$

From the hand side, instead, we obtain

$$\beta_h\Bigg\{(1-p(\omega^*)(1-\beta)\beta^{N-1}\Big((1-\alpha)^{2N+1}\tilde{p}^2+(1-\alpha)^N\beta^{N+1}\tilde{p}(1-\tilde{p})+$$

$$+(1-\alpha)^{N-1}\beta^N\tilde{p}(1-\tilde{p})+\beta^{2N+1}(1-\tilde{p})^2\Big)+$$

$$+(1-\tilde{p})\beta^N\Big((1-\alpha)^{2N+1}p(\omega^*)\tilde{p}+(1-\alpha)^{N+1}\beta^N p(\omega^*)(1-\tilde{p})+$$

$$+(1-\alpha)^N\beta^{N+1}\tilde{p}(1-p(\omega^*))+\beta^{2N+1}(1-p(\omega^*))(1-\tilde{p})\Big)+$$

$$-(1-\tilde{p})\beta^{N-1}\Big((1-\alpha)^{2N+1}p(\omega^*)\tilde{p}+(1-\alpha)^N\beta^{N+1}p(\omega^*)(1-\tilde{p})+$$

$$+(1-\alpha)^{N+1}\beta^N\tilde{p}(1-p(\omega^*))+\beta^{2N+1}(1-p(\omega^*))(1-\tilde{p})\Big)\Bigg\}=$$

$$=\beta_h\Bigg\{(1-\alpha)^{2N+1}(1-\beta)\beta^{N-1}(1-p(\omega^*))\tilde{p}^2+(1-\alpha)^N(1-\beta)\beta^{2N}(1-p(\omega^*))\tilde{p}(1-\tilde{p})+$$

$$+(1-\alpha)^{N+1}(1-\beta)\beta^{2N-1}(1-p(\omega^*))\tilde{p}(1-\tilde{p})+$$

$$-(1-\alpha)^{2N+1}(1-\beta)\beta^{N-1}p(\omega^*)\tilde{p}(1-\tilde{p})+$$

$$-(1-\alpha)^N\beta^{2N}(1-\tilde{p})\Big(p(\omega^*)(1-\tilde{p})-\beta\tilde{p}(1-p(\omega^*))\Big)+$$

$$-(1-\alpha)^{N+1}\beta^{2N-1}(1-\tilde{p})\Big((1-p(\omega^*))\tilde{p}-\beta p(\omega^*)(1-\tilde{p})\Big)\Bigg\}=$$

$$=\beta_h(1-\alpha)\Bigg\{(1-\alpha)^{2N}(1-\beta)\beta^{N-1}\Big(\tilde{p}^2-p(\omega^*)\tilde{p}^2-p(\omega^*)\tilde{p}+p(\omega^*)\tilde{p}^2\Big)+$$

$$+(1-\alpha)^{N-1}\beta^{2N}(1-\tilde{p})\Big((1-\beta)(\tilde{p}-p(\omega^*)\tilde{p})-p(\omega^*)+p(\omega^*)\tilde{p}+\beta\tilde{p}-\beta p(\omega^*)\tilde{p}\Big)+$$

$$+(1-\alpha)^N\beta^{2N-1}(1-\tilde{p})\Big((1-\beta)(\tilde{p}-p(\omega^*)\tilde{p})-\tilde{p}+p(\omega^*)\tilde{p}+\beta p(\omega^*)-\beta p(\omega^*)\tilde{p}\Big)\Bigg\}=$$

$$=\beta_h(1-\alpha)\Bigg\{(1-\alpha)^{2N}(1-\beta)\beta^{N-1}\tilde{p}(\tilde{p}-p(\omega^*))+$$

$$+(1-\alpha)^{N-1}\beta^{2N}(1-\tilde{p})(\tilde{p}-p(\omega^*))+$$

$$+(1-\alpha)^N\beta^{2N}(1-\tilde{p})(p(\omega^*)-\tilde{p})\Bigg\}=$$

$$=\beta_h(1-\alpha)(\tilde{p}-p(\omega^*))\Bigg\{(1-\alpha)^{2N}\beta^{N-1}(1-\beta)\tilde{p}+(1-\alpha)^{N-1}\beta^{2N}(1-\tilde{p})\alpha\Bigg\}.$$

As $(\tilde{p} - p(\omega^*))$ and the term in braces are non negative and common between the left and the right hand side of the original inequality, $E(\gamma_N \mid \omega^*) > E(\gamma_{N+1} \mid \omega^*)$ if and only if $(1-\alpha_h)\beta_l > (1-\alpha_l)\beta_h$. In this case, $E(\gamma_N \mid \omega^*) > E(\gamma_{N+1} \mid \omega^*) > \rho$, meaning that the expected reputation of the marginal individual, higher than the prior, lowers as the number of agents, N, increases. The opposite holds whenever $(1-\alpha_h)\beta_l < (1-\alpha_l)\beta_h$. ■

Additional beliefs with N agents with conditionally independent signals ω

Before starting with the analysis, we need to introduce some additional notation. First, we define $\tilde{\tilde{p}} \equiv \int_{\underline{\omega}}^{\omega^*} \mathbb{P}(x = g \,|\, \omega) \frac{dF(\omega)}{F(\omega^*)}$, that is the market belief on the state of the world being good after observing a manager choosing a safe project. We also use the definition $\mu \equiv \mathbb{P}(x = g)$, that is the unconditional probability of the state of the world being good.

We report below the steps to calculate the belief about the ability of the N^{th} manager, following the abandonment of the risky project after the experiment, when the market cannot assess the state of the world. There are N such beliefs and we only report the one where all the other $t-1$ managers choosing the risky project discard it following the experiment, while the remaining $N-t$ choose the safe one.

$$\gamma_N^{0t} \equiv \mathbb{P}(\theta_N = h \,|\, s_{N,\ldots,N-(t-1)} = b, \omega_{N,\ldots,N-(t-1)} > \omega^*, \omega_{1,\ldots,N-t} < \omega^*) =$$

$$= \frac{\mathbb{P}(\theta_N = h, s_{N,\ldots,N-(t-1)} = b, \omega_{N,\ldots,N-(t-1)} > \omega^*, \omega_{1,\ldots,N-t} < \omega^*)}{\mathbb{P}(s_{N,\ldots,N-(t-1)} = b, \omega_{N,\ldots,N-(t-1)} > \omega^*, \omega_{1,\ldots,N-t} < \omega^*)}$$

In evaluating $\mathbb{P}(s_{N,\ldots,N-(t-1)} = b, \omega_{N,\ldots,N-(t-1)} > \omega^*, \omega_{1,\ldots,N-t} < \omega^*)$, notice that it equals

$$\mathbb{P}(s_{N,\ldots,N-(t-1)} = b, \omega_{N,\ldots,N-(t-1)} > \omega^*, \omega_{1,\ldots,N-t} < \omega^*, x = g) +$$

$$+\mathbb{P}(s_{N,\ldots,N-(t-1)} = b, \omega_{N,\ldots,N-(t-1)} > \omega^*, \omega_{1,\ldots,N-t} < \omega^*, x = b)$$

Furthermore, we can compute $\mathbb{P}(s_{N,\ldots,N-(t-1)} = b, \omega_{N,\ldots,N-(t-1)} > \omega^*, \omega_{1,\ldots,N-t} <$

$\omega^*, x = g)$ as

$$\mathbb{P}(s_N = b \mid s_{N-1,\dots,N-(t-1)} = b, \omega_{N,\dots,N-(t-1)} > \omega^*, \omega_{1,\dots,N-t} < \omega^*, x = g) \times$$

$$\times \mathbb{P}(s_{N-1,\dots,N-(t-1)} = b, \omega_{N,\dots,N-(t-1)} > \omega^*, \omega_{1,\dots,N-t} < \omega^*, x = g)$$

Since the realizations of the experiment are conditionally independent, $\mathbb{P}(s_N = b \mid s_{N-1,\dots,N-(t-1)} = b, \omega_{N,\dots,N-(t-1)} > \omega^*, \omega_{1,\dots,N-t} < \omega^*, x = g) = \mathbb{P}(s_N = b \mid x = g) = (1-\alpha)$. Thus, reiterating the procedure, and then using the conditionally independence of signals ω, the last expression becomes

$$(1-\alpha)^t \mathbb{P}(\omega_{N,\dots,N-(t-1)} > \omega^*, \omega_{1,\dots,N-t} < \omega^*, x = g) =$$

$$= (1-\alpha)^t \mathbb{P}(\omega_N > \omega^* \mid \omega_{N-1,\dots,N-(t-1)} > \omega^*, \omega_{1,\dots,N-t} < \omega^*, x = g) \times$$

$$\times \mathbb{P}(\omega_{N-1,\dots,N-(t-1)} > \omega^*, \omega_{1,\dots,N-t} < \omega^*, x = g) =$$

$$(1-\alpha)^t \frac{\tilde{p}^t \tilde{\tilde{p}}^{N-t}}{\mu^{N-1}} (1 - F(\omega^*))^t F(\omega^*)^{N-t}$$

Computing each component of the belief in a similar way and simplifying the term $(1 - F(\omega^*))^t F(\omega^*)^{N-t}$, we get that

$$\gamma_N^{0t} = \rho \, \frac{(1-\alpha_h)(1-\alpha)^{t-1} \frac{\tilde{p}^t \tilde{\tilde{p}}^{N-t}}{\mu^{N-1}} + \beta_h \beta^{t-1} \frac{(1-\tilde{p})^t (1-\tilde{\tilde{p}})^{N-t}}{(1-\mu)^{N-1}}}{(1-\alpha)^t \frac{\tilde{p}^t \tilde{\tilde{p}}^{N-t}}{\mu^{N-1}} + \beta^t \frac{(1-\tilde{p})^t (1-\tilde{\tilde{p}})^{N-t}}{(1-\mu)^{N-1}}}$$

Results with N agents with conditionally independent signals

ω

Notice first that Lemma 1 holds also in this economy. With some abuse of notation, we define

$$\gamma_N^{0t}(\omega_1^*) = \rho \frac{(1-\alpha_h)(1-\alpha)^{t-1}\frac{p(\omega_1^*)\tilde{p}^{t-1}\tilde{\tilde{p}}^{N-t}}{\mu^{N-1}} + \beta_h \beta^{t-1}\frac{(1-p(\omega_1^*))(1-\tilde{p})^{t-1}(1-\tilde{\tilde{p}})^{N-t}}{(1-\mu)^{N-1}}}{(1-\alpha)^t\frac{p(\omega_1^*)\tilde{p}^{t-1}\tilde{\tilde{p}}^{N-t}}{\mu^{N-1}} + \beta^t\frac{(1-p(\omega_1^*))(1-\tilde{p})^{t-1}(1-\tilde{\tilde{p}})^{N-t}}{(1-\mu)^{N-1}}}$$

and

$$\gamma_N^{0t}(\omega_2^*) = \rho \frac{(1-\alpha_h)(1-\alpha)^{t-1}\frac{p(\omega_2^*)\tilde{p}^{t-1}\tilde{\tilde{p}}^{N-t}}{\mu^{N-1}} + \beta_h \beta^{t-1}\frac{(1-p(\omega_2^*))(1-\tilde{p})^{t-1}(1-\tilde{\tilde{p}})^{N-t}}{(1-\mu)^{N-1}}}{(1-\alpha)^t\frac{p(\omega_2^*)\tilde{p}^{t-1}\tilde{\tilde{p}}^{N-t}}{\mu^{N-1}} + \beta^t\frac{(1-p(\omega_2^*))(1-\tilde{p})^{t-1}(1-\tilde{\tilde{p}})^{N-t}}{(1-\mu)^{N-1}}}$$

where $\omega_1^* > \omega_2^*$. Following the same procedure that we applied in the proof of Proposition 3 it is easy to see that $\gamma_N^{0t}(\omega_1^*) > \gamma_N^{0t}(\omega_2^*)$ if and only if $\frac{1-\alpha_h}{\beta_h} > \frac{1-\alpha}{\beta}$.

As when observing a manager choosing the risky project, the market is more optimistic than the marginal manager about the state of the world ($\tilde{p} > p(\omega^*)$), this implies that

$$\gamma_N^{0t}(\tilde{\omega}) = \rho \frac{(1-\alpha_h)(1-\alpha)^{t-1}\frac{\tilde{p}^t\tilde{\tilde{p}}^{N-t}}{\mu^{N-1}} + \beta_h \beta^{t-1}\frac{(1-\tilde{p})^t(1-\tilde{\tilde{p}})^{N-t}}{(1-\mu)^{N-1}}}{(1-\alpha)^t\frac{\tilde{p}^t\tilde{\tilde{p}}^{N-t}}{\mu^{N-1}} + \beta^t\frac{(1-\tilde{p})^t(1-\tilde{\tilde{p}})^{N-t}}{(1-\mu)^{N-1}}}$$

$$>$$

$$\rho \frac{(1-\alpha_h)(1-\alpha)^{t-1}\frac{p(\omega^*)\tilde{p}^{t-1}\tilde{\tilde{p}}^{N-t}}{\mu^{N-1}} + \beta_h \beta^{t-1}\frac{(1-p(\omega^*))(1-\tilde{p})^{t-1}(1-\tilde{\tilde{p}})^{N-t}}{(1-\mu)^{N-1}}}{(1-\alpha)^t\frac{p(\omega^*)\tilde{p}^{t-1}\tilde{\tilde{p}}^{N-t}}{\mu^{N-1}} + \beta^t\frac{(1-p(\omega^*))(1-\tilde{p})^{t-1}(1-\tilde{\tilde{p}})^{N-t}}{(1-\mu)^{N-1}}} = \gamma_N^{0t}(\omega^*)$$

for any t whenever $\frac{1-\alpha_h}{\beta_h} > \frac{1-\alpha}{\beta}$. This, in turn, implies that Lemma 2 holds as well as before, and therefore that Proposition 3 applies also to this economy.

In order to obtain the expected reputation of the marginal individual from choosing the risky project, we need to calculate the weight he assigns to the posterior γ_N^{0t} for any t. We then show that such weights goes to zero as the number of managers tends to infinity. This implies that the difference between the expected reputation, obtained weighting the market's posteriors, and the expected self-assessment of the manager, obtained using the same weights but different posteriors, vanishes in the limit. Indeed, the only posteriors in which the market and the manager would disagree upon has weights approaching zero.

Consider the Nth manager, that receives the signal $\omega = \omega^*$. In his expected reputation, the weight associated to the event in which some specific $N-t$ managers choose the safe project, while the other $t-1$ managers and himself discard the risky one after the experiment is given by

$$\mathbb{P}(s_{N,\dots,N-(t-1)} = b, \omega_{N-1,\dots,N-(t-1)} > \omega^*, \omega_{1,\dots,N-t} < \omega^* \mid \omega_N = \omega^*) =$$

$$\mathbb{P}(s_{N,\dots,N-(t-1)} = b, \omega_{N-1,\dots,N-(t-1)} > \omega^*, \omega_{1,\dots,N-t} < \omega^* \mid \omega_N = \omega^*, x = g)\times$$

$$\times\mathbb{P}(x = g \mid \omega_N = \omega^*)+$$

$$\mathbb{P}(s_{N,\dots,N-(t-1)} = b, \omega_{N-1,\dots,N-(t-1)} > \omega^*, \omega_{1,\dots,N-t} < \omega^* \mid \omega_N = \omega^*, x = b)\times$$

$$\times\mathbb{P}(x = b \mid \omega_N = \omega^*)$$

Notice that

$$\mathbb{P}(s_{N,\ldots,N-(t-1)} = b, \omega_{N-1,\ldots,N-(t-1)} > \omega^*, \omega_{1,\ldots,N-t} < \omega^* \mid \omega_N = \omega^*, x = g) =$$

$$= \mathbb{P}(s_N \mid s_{N-1,\ldots,N-(t-1)} = b, \omega_{N-1,\ldots,N-(t-1)} > \omega^*, \omega_{1,\ldots,N-t} < \omega^*, \omega_N = \omega^*, x = g)\times$$

$$\times\mathbb{P}(s_{N-1,\ldots,N-(t-1)} = b, \omega_{N-1,\ldots,N-(t-1)} > \omega^*, \omega_{1,\ldots,N-t} < \omega^* \mid \omega_N = \omega^*, x = g)$$

Using conditionally independence of the signals ω and the signals from the experiment, this equals

$$(1-\alpha)^t \frac{\tilde{p}^{t-1}\tilde{\tilde{p}}^{N-t}}{\mu^{N-1}}(1-F(\omega^*))^{t-1}F(\omega^*)^{N-t}$$

Notice now that there are $\binom{N-1}{t-1}$ possible ways in which some specific $N-t$ managers choose the safe project, while the other $t-1$ managers and the N^{th} manager discard the risky one after the experiment. Indeed these are the ways in which one could select $t-1$ managers out of $N-1$ (or, that is equivalent, $N-t$ managers out of $N-1$). So the weight associated to the posterior γ_N^{0t} is given by

$$\binom{N-1}{t-1}F(\omega^*)^{N-t}(1-F(\omega^*))^{t-1}\times$$

$$\times\left[(1-\alpha)^t\left(\frac{\tilde{p}}{\mu}\right)^{t-1}\left(\frac{\tilde{\tilde{p}}}{\mu}\right)^{N-t}p(\omega) + \beta^t\left(\frac{1-\tilde{p}}{1-\mu}\right)^{t-1}\left(\frac{1-\tilde{\tilde{p}}}{1-\mu}\right)^{N-t}(1-p(\omega))\right] =$$

$$= (N-1)...(N-t+1)\left[F(\omega^*)\frac{\tilde{\tilde{p}}}{\mu}\right]^{N-t} \underbrace{\frac{1}{(t-1)!}(1-F(\omega^*))^{t-1}\left[(1-\alpha)^t\left(\frac{\tilde{p}}{\mu}\right)^{t-1}p(\omega^*)\right]}_{\equiv A_t} +$$

$$+(N-1)...(N-t+1)\left[F(\omega^*)\frac{1-\tilde{\tilde{p}}}{1-\mu}\right]^{N-t} \underbrace{\frac{1}{(t-1)!}(1-F(\omega^*))^{t-1}\left[\beta^t\left(\frac{1-\tilde{p}}{1-\mu}\right)^{t-1}(1-p(\omega^*))\right]}_{\equiv B_t}$$

Consider $a_N \equiv (N-1)...(N-t+1)\left[F(\omega^*)\frac{\tilde{\tilde{p}}}{\mu}\right]^{N-t} A_t$, first. As N goes to infinity, this expression approaches $(N-1)^{t-1}\left[F(\omega^*)\frac{\tilde{\tilde{p}}}{\mu}\right]^{N-t} A_t$. Consider now $\frac{a_{N+1}}{a_N} = \left(\frac{N}{N-1}\right)^{t-1} F(\omega^*)\frac{\tilde{\tilde{p}}}{\mu}$. Using the definitions $\tilde{\tilde{p}} \equiv \int_{\underline{\omega}}^{\omega^*} \mathbb{P}(x=g\,|\,\omega)\frac{dF(\omega)}{F(\omega^*)}$ and $\mu \equiv \mathbb{P}(x=g)$, it follows that $F(\omega^*)\frac{\tilde{\tilde{p}}}{\mu} = \frac{\int_{\underline{\omega}}^{\omega^*}\mathbb{P}(x=g\,|\,\omega)dF(\omega)}{\int_{\underline{\omega}}^{\overline{\omega}}\mathbb{P}(x=g\,|\,\omega)dF(\omega)} < 1$ since $\overline{\omega} > \omega^*$. Thus, for N large enough

$$\frac{a_{N+1}}{a_N} = \left(\frac{N}{N-1}\right)^{t-1} F(\omega^*)\frac{\tilde{\tilde{p}}}{\mu} < 1$$

Using the *ratio test*, this implies that

$$\sum_{N=t}^{\infty}(N-1)^{t-1}\left[F(\omega^*)\frac{\tilde{\tilde{p}}}{\mu}\right]^{N-t} A_t$$

converges absolutely and, thus, that

$$\lim_{N\to\infty}(N-1)^{t-1}\left[F(\omega^*)\frac{\tilde{\tilde{p}}}{\mu}\right]^{N-t} A_t = \lim_{N\to\infty}(N-1)...(N-t+1)\left[F(\omega^*)\frac{\tilde{\tilde{p}}}{\mu}\right]^{N-t} A_t = 0$$

.

Applying the same reasoning to the expression $(N-1)...(N-t+1)\left[F(\omega^*)\frac{1-\tilde{\tilde{p}}}{1-\mu}\right]^{N-t} B_t$, we obtain that $\frac{b_{N+1}}{b_N} = \left(\frac{N}{N-1}\right)^{t-1} F(\omega^*) \frac{1-\tilde{\tilde{p}}}{1-\mu} = \left(\frac{N}{N-1}\right)^{t-1} \frac{\int_{\underline{\omega}}^{\omega^*} \mathbb{P}(x=b \,|\, \omega) dF(\omega)}{\int_{\underline{\omega}}^{\overline{\omega}} \mathbb{P}(x=b \,|\, \omega) dF(\omega)} < 1$, as $\overline{\omega} > \omega^*$. Using again the *ratio test*, we thus conclude that the weights associated to each posterior γ_N^{0t} are zero in the limit, for any t. ■

Bibliography

Yann Algan, Christian Dustmann, Albrecht Glitz, and Alan Manning. The economic situation of first and second-generation immigrants in france, germany and the united kingdom. *The Economic Journal*, 120(542), 2010.

Joshua D Angrist and Jörn-Steffen Pischke. *Mostly harmless econometrics: An empiricist's companion*. Princeton university press, 2008.

Esteban Aucejo, Jonathan James, et al. The path to college education: Are verbal skills more important than math skills? Technical report, 2016.

Sukanya Basu. Age-of-arrival effects on the education of immigrant children: A sibling study. 2016.

Alberto Bisin and Thierry Verdier. The economics of cultural transmission and socialization. In *Handbook of social economics*, volume 1, pages 339–416. Elsevier, 2011.

Sandra E Black, Paul J Devereux, and Kjell G Salvanes. The more the merrier? the effect of family size and birth order on children's education. *The Quarterly Journal of Economics*, 120(2):669–700, 2005.

Hoyt Bleakley and Aimee Chin. Language skills and earnings: Evidence from childhood immigrants. *Review of Economics and Statistics*, 86(2):481–496, 2004.

Anders Böhlmark. Age at immigration and school performance: A siblings analysis using swedish register data. *Labour Economics*, 15(6):1366–1387, 2008.

George J Borjas. Assimilation, changes in cohort quality, and the earnings of immigrants. *Journal of labor Economics*, 3(4):463–489, 1985.

George J Borjas. The intergenerational mobility of immigrants. *Journal of Labor Economics*, 11(1, Part 1):113–135, 1993.

Stephen V Cameron and James J Heckman. Life cycle schooling and dynamic selection bias: Models and evidence for five cohorts of american males. *Journal of Political Economy*, 106(2):262–333, 1998.

Stephen V Cameron and James J Heckman. The dynamics of educational attainment for black, hispanic, and white males. *Journal of Political Economy*, 109(3):455–499, 2001.

Ying Chen. Career concerns and excessive risk taking. *Journal of Economics & Management Strategy*, 24(1):110–130, 2015.

Barry R Chiswick. Sons of immigrants: Are they at an earnings disadvantage? *The American Economic Review*, 67(1):376–380, 1977.

Barry R Chiswick and Noyna DebBurman. Educational attainment: analysis by immigrant generation. *Economics of Education Review*, 23(4):361–379, 2004.

Flavio Cunha and James Heckman. The technology of skill formation. Technical report, National Bureau of Economic Research, 2007.

Flavio Cunha, James J Heckman, and Susanne M Schennach. Estimating the technology of cognitive and noncognitive skill formation. *Econometrica*, 78(3):883–931, 2010.

Christian Dustmann, Tommaso Frattini, and Gianandrea Lanzara. Educational achievement of second-generation immigrants: an international comparison. *Economic Policy*, 27(69):143–185, 2012.

Michael Ewens, Ramana Nanda, and Matthew Rhodes-Kropf. Cost of experimentation and the evolution of venture capital. *Journal of Financial Economics (forthcoming)*, 2016.

Arturo Gonzalez. The education and wages of immigrant children: the impact of age at arrival. *Economics of Education Review*, 22(2):203–212, 2003.

Robert E Hall and Susan E Woodward. The burden of the nondiversifiable risk of entrepreneurship. *American Economic Review*, 100(3):1163–94, 2010.

James J Heckman and Yona Rubinstein. The importance of noncognitive skills: Lessons from the ged testing program. *The American Economic Review*, 91(2): 145–149, 2001.

James J Heckman, Jora Stixrud, and Sergio Urzua. The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics*, 24(3):411–482, 2006.

Yael V Hochberg, Alexander Ljungqvist, and Annette Vissing-Jørgensen. Informational holdup and performance persistence in venture capital. *The Review of Financial Studies*, 27(1):102–152, 2013.

Bengt Holmström. Managerial incentive problems: A dynamic perspective. *The Review of Economic Studies*, 66(1):169–182, 1999.

Ingo E Isphording, Marc Piopiunik, and Núria Rodríguez-Planas. Speaking in numbers: The effect of reading performance on math performance among immigrants. *Economics Letters*, 139:52–56, 2016.

Jacqueline S Johnson and Elissa L Newport. Critical period effects in second language learning: The influence of maturational state on the acquisition of english as a second language. *Cognitive psychology*, 21(1):60–99, 1989.

Steven N Kaplan and Antoinette Schoar. Private equity performance: Returns, persistence, and capital flows. *The Journal of Finance*, 60(4):1791–1823, 2005.

William R Kerr, Ramana Nanda, and Matthew Rhodes-Kropf. Entrepreneurship as experimentation. *The Journal of Economic Perspectives*, 28(3):25–48, 2014.

Arthur Korteweg and Morten Sorensen. Skill and luck in private equity performance. *Journal of Financial Economics*, 124(3):535–562, 2017.

Dominique Lemmermann and Regina T Riphahn. The causal effect of age at migration on youth educational attainment. *Economics of Education Review*, 2017.

Eric H Lenneberg. The biological foundations of language. *Hospital Practice*, 2(12): 59–67, 1967.

Wei Li. Changing one's mind when the facts change: incentives of experts and the design of reporting protocols. *The Review of Economic Studies*, 74(4):1175–1194, 2007.

Robert Marquez, Vikram Nanda, and M Deniz Yavuz. Private equity fund returns and performance persistence. *Review of Finance*, page rfu045, 2014.

Ramana Nanda and Matthew Rhodes-Kropf. Investment cycles and startup innovation. *Journal of Financial Economics*, 110(2):403–418, 2013.

Asako Ohinata and Jan C van Ours. Young immigrant children and their educational attainment. *Economics Letters*, 116(3):288–290, 2012.

Regina T Riphahn. Cohort effects in the educational attainment of second generation immigrants in germany: an analysis of census data. *Journal of Population Economics*, 16(4):711–737, 2003.

Stephen J Trejo. Why do mexican americans earn low wages? *Journal of Political Economy*, 105(6):1235–1268, 1997.

James Truslow Adams. The epic of america. *Boston, MA: Little, Brown & Company*, 1931.

Jan C Van Ours and Justus Veenman. Age at immigration and educational attainment of young immigrants. *Economics Letters*, 90(3):310–316, 2006.

Karen A Woodrow-Lafield, Xiaohe Xu, Thomas Kersen, and Bunnak Poch. Naturalization of us immigrants: Highlights from ten countries. *Population Research and Policy Review*, 23(3):187–218, 2004.