UAB
Universitat Autònoma de Barcelona

ESSAYS ON FAMILY ECONOMICS

Jacek Dominik Barszczewski

UAB
Universitat Autònoma de Barcelona

ESSAYS ON FAMILY ECONOMICS

Author: Jacek Dominik Barszczewski

Advisor: Dr Joan Llull

June 2023

ACKNOWLEDGEMENTS

Words cannot adequately convey the depth of my gratitude to my advisor, Joan Llull. I am forever indebted to him for encouraging me to embark on the challenging journey of pursuing a PhD and never leaving me alone throughout this incredible adventure. Joan was with me throughout all the ups and downs on this bumpy road. He relentlessly provided me with constant and unconditional support. Joan's exceptional scientific guidance was essential in shaping my economic and quantitative skills. His unwavering belief in my abilities was an unending source of motivation and inspiration, empowering me to reach new heights in my academic journey. Moreover, I am immensely grateful for Joan's infectious sense of humor, remarkable patience, and boundless enthusiasm, which infused our interactions with joy and fostered an environment of fruitful collaboration.

Many great researchers helped me during the course of the PhD and contributed to its final outcome. I am deeply grateful to Nezih Guner for his invaluable contributions to my research. His insightful comments, combined with his exceptional sense of humor, greatly enriched my work. I also extend my sincere thanks to Hanna Wang for her great support and long discussions that significantly enhanced the quality of my research. I would like to express my gratitude to all the presenters and participants of the Applied Seminars and researchers contributing to the IDEA program. Their expertise and feedback played a vital role in shaping my research and broadening my understanding of the field.

Thanks also go to George-Levi Gayle and Limor Golan for hosting me at Washington University in St. Louis. Their exceptional guidance and highly valuable feedback on my research significantly contributed to its quality. Moreover, they also provided me with invaluable opportunities to collaborate with researchers both at Washington University and the Federal Reserve Bank of St. Louis. I would also like to express my sincere appreciation to my esteemed co-author, Prasanthi Ramakrishnan, whom I had the privilege of meeting during my stay in St. Louis. Our collaboration was tremendously beneficial, and I greatly benefited from her expertise and insights.

Special thanks to Inés Macho-Stadler, David Pérez-Castrillo, and the entire IDEA department for their support and guidance throughout my PhD, particularly during the job market phase. Their expertise and mentorship have been invaluable to me. I also extend my sincere thanks to the department administration, Àngels López-Garcia and Mercé Vincente, for their efficiency, responsiveness, and kindness in handling all administrative matters. Their assistance made the administrative process smooth and effortless, and their support was relentless throughout my time in the department.

I am also grateful to all my colleagues at IDEA, including visiting students, for the memorable moments and stimulating conversations we shared together. I especially would like to thank the members of the Equality group for their insightful comments and engaging discussions. Additionally, I want to express my deep appreciation to Karol Madoń, with whom I embarked on this intellectual adventure at the Warsaw School of Economics. His unwavering support throughout my entire academic journey, including the PhD time, was invaluable to me.

I would like to express my sincere appreciation to my parents, Jadwiga and Krzysztof, for their continuous support and the countless efforts they invested in raising me. Their guidance and care have played a pivotal role in shaping my character and fueling my aspirations. I am also deeply grateful to my sister Natalia, whose unwavering belief in me and steadfast presence have been a constant source of encouragement. To my dear friends in Warsaw, your enduring friendship and support have meant a lot to me despite the physical distance that separates us. A special thanks goes to my friends from Gimnazjum 81, whose witty yet caring comments have always kept me grounded.

Last but not least, I want to express my deepest gratitude to my partner, Katherina Thomas. Throughout my entire PhD journey, she has been my enduring source of joy and a guiding light during the darkest moments. Her presence, genuine honesty, and invaluable advice were instrumental in my personal and academic growth. Her relentless belief in my potential help me to reach new heights. I am truly fortunate to have her by my side, and I am forever grateful for her unwavering support.

I gratefully acknowledge financial support from the FPI Fellowship of the Ministerio de Ciencia, Innovación y Universidades linked to the project SEV-2015-0563-19-1, the European Research Council (ERC) through Starting Grant no. 804989, and the Generalitat de Catalunya through the research grant 2021 SGR 00194.

CONTENTS

INTRODUCTION

In this thesis, I employ empirical methods to address research questions related to marriage and family behavior. In the first chapter, I investigate the influence of labor market conditions for immigrants on their location and marriage choices. These two decisions are crucial for immigrants' social integration in the host country. In particular, I concentrate on changes in marriage and location patterns resulting from labor market integration policies. In the second chapter (joint with Katherina Thomas), we examine the role of parents in the process of children's human capital accumulation. We analyze how parenting style affects a child's cognitive and non-cognitive skills. In the third chapter (joint with Prasanthi Ramakrishnan), we focus on family formation, when two partners live far apart, requiring the migration of one partner from their parental household to a future partner's household. In that chapter, we study the implications of such long-distance marriage migration for women within-household bargaining power in India.

In Chapter 1 of this thesis, *Location Choice, Labor Market Conditions, and Marital Sorting Among Immigrants*, I analyze how measures of social integration, like the share of marriages with natives and immigrants' spatial concentration, change under different labor market integration policy scenarios. I first show correlations between immigrants' labor market outcomes, marital patterns, and spatial distribution. Then, using German data, I estimate a structural model with location, marriage, and labor supply decisions. The model reflects two trade-offs immigrants face: a) partner choice: "marry your like" vs. economic gains from marriage with a native, and b) location choice: a region with higher wages vs. a region with better marriage opportunities. Model simulations reveal that: 1) reducing the immigrant-native income gap by 25% decreases immigrants' spatial concentration (by 2.9%) but lowers the share of immigrant women married to natives (by 2 pp); 2) declining the regional wage gap by 50% significantly reduces immigrants' spatial concentration (by 15%), increases the share of immigrant men married to native (by 1.1 pp), but decreases the share of immigrant women married to natives (by 0.6 pp). I also find that ignoring adjustments in location and marriage choices under both policies overstates the decrease in immigrant-native income inequality and underpredicts the welfare gains. The reason for that is when immigrants' labor market position improves, they give up part of their income gains and marry natives less often to satisfy their taste for similarity in partners' origin, increasing their welfare.

In Chapter 2, *Parenting Style and Children's Skill Development*, we examine the influence of parenting style on cognitive and non-cognitive skill development in middle childhood and adolescence. Using Australian panel data, we estimate the effects of various dimensions of parenting style on skill development. To address identification issues, we exploit the panel structure of the data, incorporate a rich set of controls and employ multiple econometric specifications. Our findings indicate that parental hostility, characterized by a lack of praise for the child and display of anger during punishments, as well as inconsistency in enforcing rules, have a negative impact on non-cognitive skills. Explaining the implemented rules to the child has a smaller negative effect, while parental warmth exhibits a small positive effect. The estimated effects are substantial in magnitude. A one standard deviation increase in hostility is associated with a decrease in children's non-cognitive skills ranging from 0.12 to 0.35 standard deviations, depending on the econometric specification. We find a limited impact of parenting style on children's cognitive skill development overall. Our results suggest that targeting hostility in parenting skill training could substantially improve non-cognitive skills.

In Chapter 3, *Across-District Marriage Migration in India*, we examine how long-distance marriage migration contributes to within-household inequality in India. Given the significant regional skew in the sex ratio, it is common for women to move outside their district for marriage. While such migration increases the distance from their natal home, it may also be advantageous for women as they move to districts with a more imbalanced sex ratio. As it is not clear how these two mechanisms play out, we employ various empirical methods to investigate the influence of marriage migration on women's bargaining power. First, using logistic regression analysis, we find that women are more likely to migrate for marriage to regions with a more imbalanced sex ratio and to rural households where the household head possesses at least primary education. Second, we construct a static marriage market model incorporating across-district marriage migration. Through this model, we analyze the within-household bargaining power of local and migrant women. The results reveal a negative correlation between men's marriage surplus and the probability of marrying a woman from another district. This suggests that arriving women may possess higher bargaining power and benefit from moving away from their home district. Given the inherent limitations of the static marriage market model, we propose a theoretical collective household model to further examine the relationship between marriage migration and women's bargaining power.

1. LOCATION CHOICE, LABOR MARKET CONDITIONS, AND MARITAL SORTING AMONG IMMIGRANTS

1.1. Introduction

Over the last decade, more than 45 million people permanently migrated to OECD countries[1], increasing the share of the foreign-born population by 16%[2]. This growing number puts a spotlight on immigrant integration in the public debate. In response, policymakers implement various programs that support integration. The majority of these programs[3] focus on integrating immigrants into the labor market. If successful, they improve the immigrant economic situation. Despite the economic dimension of integration, there is also the social one. The intensity of immigrant interactions with natives also contributes to their level of integration. Among others, the social dimension is measured by the frequency of marriages with natives and immigrant spatial concentration (Lazear, 1999, Danzer and Yaman, 2013, Boeri et al., 2015, Cutler et al., 2008). What are the consequences of labor market integration policies for the social dimensions of integration? How do the immigrants' marital patterns change? Do immigrants adjust their location choices? To what extent do those changes impact income inequality and welfare?

The answer to these questions is not trivial and depends on several factors. While searching for a partner, immigrants face the decision to marry another immigrant or a native. On the one hand, in most OECD countries, foreign-born earn less, on average, than natives. In this sense, intermarriage[4] may improve their financial situation. On the other hand, immigrants, likewise natives, show preferences for similarity, which makes other immigrants more attractive. Consequently, in one scenario, improving immigrants' labor market outcomes raises their attractiveness in native eyes, increases the intermarriage rate, and fosters immigrant social integration. In the alternative scenario, it decreases relative gain from cohabitation with a native, leads to more marriage between immigrants, and mitigates the positive effect of labor market policy. Moreover, marital patterns depend on the partners' availability, which, in turn, depends on location choices. Immigrants can trade regions with higher wages for regions with more immigrant partners, assuming they have a taste for similarity. As a result, depending on the character of changes in the labor market and preferences for similarity, immigrants might adjust location choices in a way that leads to a decrease or increase in their spatial concentration. Further, the direction of changes in location and marriage patterns impact household resources and thus affect income inequality and welfare.

[1] Author's calculation for years 2010-2019 based on *International Migration Outlook 2021*

[2] Ibid.

[3] The most common integration programs are active labor market integration policies, i.e., language training, labor market training and work practice, subsidized employment, and job search assistance.

[4] The existing literature does not uniquely define the *intermarriage* term. In general, *intermarriage*

To analyze potential consequences of labor market integration policies and capture the relevant trade-offs, I build a structural model with an equilibrium marriage market in which immigrants and natives choose their location, find partners, and optimize their labor supply. I estimate the model with German microdata and quantify the effect of labor market policy outcomes on marriage and location patterns. Further, I conduct welfare and income inequality analyses to understand how controlling for adjustments in marriage and region choices changes the initial economic effect of the labor market policies. I propose the modeling approach that allows for answering the research questions and conducting relevant analyses in three ways. First, location choices depend on labor market conditions so that I can predict the spatial concentration of immigrants under different labor market policy scenarios. Second, while choosing a partner, agents take into account future household income. Due to that, I can simulate how different labor market conditions change marriage choices in equilibrium. And third, in my policy exercises, I can carefully control for interdependence between location and marriage choices.

The model presented in this paper builds on the recent works on the matching models by Chiappori and Mazzocco (2017), Adda et al. (2020), Galichon and Salanié (2021) in the spirit of Becker (1973, 1974). Agents make a labor supply decision within the static collective household framework. Natives and immigrants differ in wages and leisure preferences by education (college vs. noncollege) and region (North, South and West)[5] to capture the observed variation in income and labor supply. I allow wages to vary by marriage status and by partner's origin. By that, I account for potential immigrants' wage premium from intermarriage empirically shown by, i.e., Meng and Gregory (2005), Basu (2015), Elwert and Tegunimataka (2016). Further, the marriage surplus depends on the future household budget, which generates differences in marital gains by the partner's education and origin. Finally, agents have preferences toward similarity in origin and education to capture the observed assortative mating in marriage patterns.

Natives and immigrants make lifetime location choices based on regional characteristics, such as expectations towards marriage and labor market outcomes and the value of local amenities. Since locations differ in the level of wages by origin, education, and gender, agents have incentives to distribute disproportionally across regions. As a result, the underrepresented types have more bargaining power in the local marriage market and

refers to marriage outside own social group. It has traditionally been restricted only to actual formal marriage. Nowadays, this way of defining intermarriages seems to omit the other common forms of partnership. Possibly due to social pressure, immigrant-native couples even more often avoid a formal framework (Benson, 1981). It stresses the need to extend the *intermarriage* definition to other forms of partnership. The other problem emerges with the definition of own social group. (see Rodríguez-García (2015) and Elwert (2018) for further discussion). This paper uses the term *intermarriage* as an informal and formal partnership of foreign-born and native-born individuals. By *nonintermarriage*, I denote any other form of an informal and formal partnership.

[5]Following the Federal Statistical Office of Germany, I define four macro-regions: South - Hesse, Baden-Wuerttemberg, and Bavaria; West - North-Rhine-Westfalia, Rhineland-Palatinate, Saarland; North - Schleswig-Holstein, Hamburg, Lower-Saxony, Bremen and Berlin. East region is dropped from analysis due to a very small migration population.

benefit from the higher transfers in the matching process (I model the marriage market in a frictionless framework with transferable utility). The transfer sizes impact the expected utility of settling in a region and influence its attractiveness in equilibrium, directly linking location and marriage choices. This link is especially crucial for immigrants since their number is relatively small, so any changes in spatial distribution have a profound impact on marriage outcomes (van Tubergen and Maas, 2007, Harris and Ono, 2005, Choi and Tienda, 2017). Region's utility also includes the exogenous amenity index, which I create, following Diamond (2016) in the estimation process, based on the broad set of variables, i.e., access to public transport or number of severe crimes.

Thanks to the model structure, I conduct the estimation in three steps, starting with the household problem. It is a standard static labor supply problem, so labor market and leisure parameters are identified by observed variations in wages and labor supply choices. I fit this part of the model to the data from the German Socioeconomic Panel, waves 1984-2018. In the second step, I use consistent estimates of the labor market parameters to predict the total household economic gain. Thanks to that, I can later estimate tastes for similarity and endogenous transfers on the marriage market for the baseline scenario. The marriage market equilibrium conditions entirely determine the intrahousehold allocation of the economic gains for all possible matches. I identify the partner preference parameters by observed marital outcomes, following the approach by Choo and Siow (2006). In the last step, I use marriage parameter estimates to predict the expected utility of participating in regional marriage and labor markets. This value, together with the amenity index, determines the location choices. Next, I identify the taste for amenities by across-cohort variation in location choice probabilities. For the last two steps, I fit the model using the German Microcensus 2006, 2010, and 2015.

I find a significant gap between the earnings of immigrants and natives by gender and education. This finding is in line with the less than perfect international transferability of human capital (i.e., Chiswick and Miller (2009)). As a result, households with immigrants have lower disposable income than those with only natives of the same education level. Moreover, the estimated immigrant-native wage gap varies across regions. It means that, to some extent, the difference in the distributions of immigrants and natives across space is driven by variations in labor market outcomes. Next, the expected economic surplus generated by each type of household depends on household income and the value of leisure. Keeping the same level of partners' education, I find that the surplus is higher in the case of immigrant-native households due to higher preferences for leisure among immigrants (complementary effect). It makes mixed unions more attractive from an economic point of view.

Marriage market equilibrium conditions, preference parameters, and expected economic surplus from marriage determine agents' marital choices. Estimated similarity parameters imply that agents prefer to match with partners of the same origin and education. However, the preferences for similarity are stronger in origin than in education. I also find that estimated endogenous transfers between agents show some patterns in the bargaining power of agents in the marriage market. On average, agents with higher potential earnings and are more scarce in the population have a better negotiating position. As a result, they obtain higher endogenous transfer in the marriage market. It means that agents have incentives to choose a location with a lower wage but fewer agents of the same type and compensate for the loss in income by higher marriage market transfer. In this way, marriage market conditions partially counteract labor market motives for location choices.

Subsequently, I use the estimated model parameters to quantify the effect of labor market integration policies on intermarriage and spatial concentration of immigrants. I do so by simulating two counterfactual scenarios. In the first scenario, a government introduces a country-wide policy that reduces an immigrant-native wage gap (i.e. publicly available language courses). As a result, the increase in immigrants' wages equals 25% of the initial value of the gap by region, gender, and education. This increase is equivalent to an average wage rise of 4.6% for foreign-born men and 7.5% for foreign-born women. Under the second scenario, I assume that a government targets the regional variation in immigrants' wages (i.e. locally subsidize employment). The policy increases the earnings of non-native residents in regions with an overall lower income level. In the aftermath, the differences between the region with the highest wages and the remaining ones reduce by 50% of the initial value of the gap by gender and education level. This pay rise is equal to a 3.5% increase in the average wages of immigrants.

First, I show that outcomes of introduced policies lead to a decrease in the spatial concentration of immigrants[6]. The decline is more substantial in the case of a reduction in regional variation in wages and is equal to around 10% for noncollege- and 32% for college-educated immigrants. Therefore, to a different extent, both scenarios ease financial incentives to concentrate in the region with the most favorable labor market. It means that there is a positive impact of analyzed policies on social integration via adjustments in location choices. Next, I find that the effects of both policies on marriage patterns are mixed and vary by gender. Immigrant men are less likely to stay single (from -1.6 to -8.2 pp, stronger effect when the immigrant-native gap is reduced). Further, they are also more likely to be intermarried (from 0.9 to 2.2 pp). In the case of women, the increase in the number of marriages is smaller, and the probability of marrying native men decreases (up to -2.4 pp). As a result, the outcomes of analyzed labor market policies have a positive impact via intermarriages on the social integration of men but a negative (to a greater extent when the immigrant-native gap is reduced) in the case of women.

Finally, I conduct welfare and income inequality analyses. The policies that increase immigrants' wages mechanically reduce income inequality between immigrants and natives[7]. However, I find that ignoring the adjustments in marriage and location choices

[6] I measure the spatial concentration by the total variation distance between uniform and observed distributions

[7] I measure income inequality between immigrants and natives as a percentage difference in the per capita income by gender and education.

leads to overprediction of a decline in income inequality. In the case of reducing the immigrant-native gap and regional wage variation, the decline equals, respectively, 5% and 6%. Unlike income inequality, I show that ignoring both adjustments is associated with the underprediction of welfare gains. While reducing the immigrant-native gap, the underprediction equals 12%. In the case of the reduction in regional wage variation, it is even higher and equals 15%. It means that when immigrants' earnings rise, they give up part of their marriage economic gains by marrying natives less often. They do so to satisfy their taste for similarity in their partner's origin, increasing their welfare gains.

This paper is related to several strands of literature. Most closely related are studies on the integration of immigrants. By conducting immigrant-native income inequality and welfare analyses, I extend the literature that studies the effects of government integration programs on the economic performance of immigrants, see among others Hayfron (2001), Lochmann et al. (2019), Joona and Nekby (2012), Cohen-Goldner and Eckstein (2008, 2010). Hayfron (2001) and Lochmann et al. (2019) study the participation of immigrants in language, while Joona and Nekby (2012) evaluate whether intensive counseling and coaching improve immigrants' employment opportunities. Cohen-Goldner and Eckstein (2008), and Cohen-Goldner and Eckstein (2010) finds a positive effect of local training on wages and labor market participation. I also quantify the effect of labor market integration policies on intermarriage and spatial concentration of immigrants. It allows me to evaluate if those policies positively impact not only the economic integration of immigrants but also its social aspects. By that, I contribute to the literature that focuses on the determinants and socio-economic consequences of non-labor aspects of integration, see among others Kalmijn and van Tubergen (2006), Dribe and Lundh (2011), Chiswick and Houseworth (2011), Grossbard and Vernon (2020), Xie and Gough (2011), Min Zhou and Logan (1989).

My analysis of the marriage market builds on previous equilibrium models with a transferable utility, such as Chiappori et al. (2018) and Calvo et al. (2021). My model structure is closest to the one proposed in Chiappori et al. (2018) regarding the marriage market and household behavior. However, I focus on immigrant integration. Hence, I distinguish individuals not only by education level but also by immigrant status. Further, the first choice in my model is location decision instead of the decisions of human capital (education) investments. The paper by Calvo et al. (2021) relates to mine in that they focus on the relationship between labor and marriage markets and estimate their model using the same German data. In my model, a labor market impacts marriage patterns through changes in economic gains. Unlike, they examine how the connection between labor and the marriage market affects home production and patterns of job matching.

Combining the location choice decision with the marriage and labor markets is a novel feature of my model. On the one hand, my location choice decision model is inspired by the tradition of spatial equilibrium models initiated by Rosen (1979) and Roback (1982) and recently popularized by Diamond (2016). Unlike all these models, the equilibrium clearing in my model occurs in the marriage market. However, given the nature of my counterfactual exercises, I abstract from labor market equilibrium for tractability reasons. Unlike in these models, however, the availability of potential partners of different types plays a crucial role in my model. On the other hand, equilibrium marriage models generally focus on a single global market. However, several reduced-form papers show that marriage market outcomes differ across space and impact location choices, see among others Costa and Kahn (2000), Compton and Pollak (2007), Chiswick and Houseworth (2011). I allow for the endogenous spatial allocation of individuals on all sides of the market, which leads to changes in bargaining power, which fundamentally affect marriage market outcomes. The paper that similarly uses a setting with endogenous sorting and marriage market is Fan and Zou (2021). Contrary to them, I distinguish individuals by origin, so I can study the effects separately for immigrants and natives instead of focusing on the determinants of the spatial distribution of economic activities.

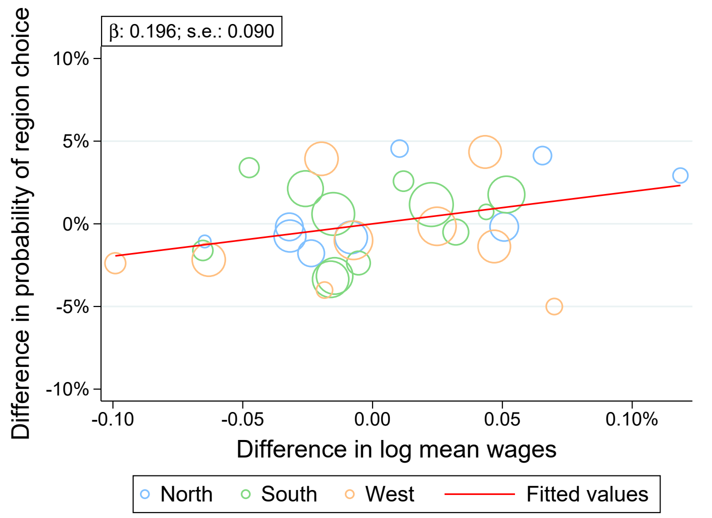
Finally, a few papers analyze the marriage patterns of immigrants and natives in an equilibrium framework. The most notable example is Adda et al. (2020). In that paper, the authors explore the trade-off between mating along cultural lines and legal status acquisition, which can positively impact labor outcomes. Adda and coauthors also study local marriage markets, but they take a geographical distribution of immigrants as given. Using my framework, I can investigate how the spatial concentration of immigrants would change given the anticipated changes in the labor market and what consequences it has for marriage patterns. The mechanisms associated with location choice are even more important while analyzing immigrants since: (a) they are more mobile than natives, so they might stronger respond to changes in location conditions; (b) they are a relatively smaller group compared to natives, so any change in the local composition of the marriage market has a more substantial impact on their marriage outcome.

The rest of the paper is organized as follows. Section 1.2 contains basic statistics and empirical facts linking location choice, marriage market, and economic integration of immigrants. Section 1.3 presents the model, while Section 1.4 discusses the data used and employed estimation strategy. Section 1.5 contains outcomes of conducted counterfactual scenarios. Section 1.6 concludes.

1.2. Data

In this section, I provide evidence linking location choice, marriage, and labor market outcomes. I use this evidence to motivate the research question and the model structure presented in the next section. To conduct the empirical analyses, I use German data. I choose to focus on Germany as it is an attractive country for immigrants from different origins. Immigrants from East European and Post-Soviet countries are the biggest immigration group, and their share in total migration stock is slightly above 30%. The second biggest group is Turkish immigrants, which share is equal to 17%. A similar share of immigrants is of Balkan origins. The last significant group is Southern European immigrants, which share is equal to 11%. The remaining 25% of immigrants came from

Figure 1.1: Difference in mean wage and spatial sorting



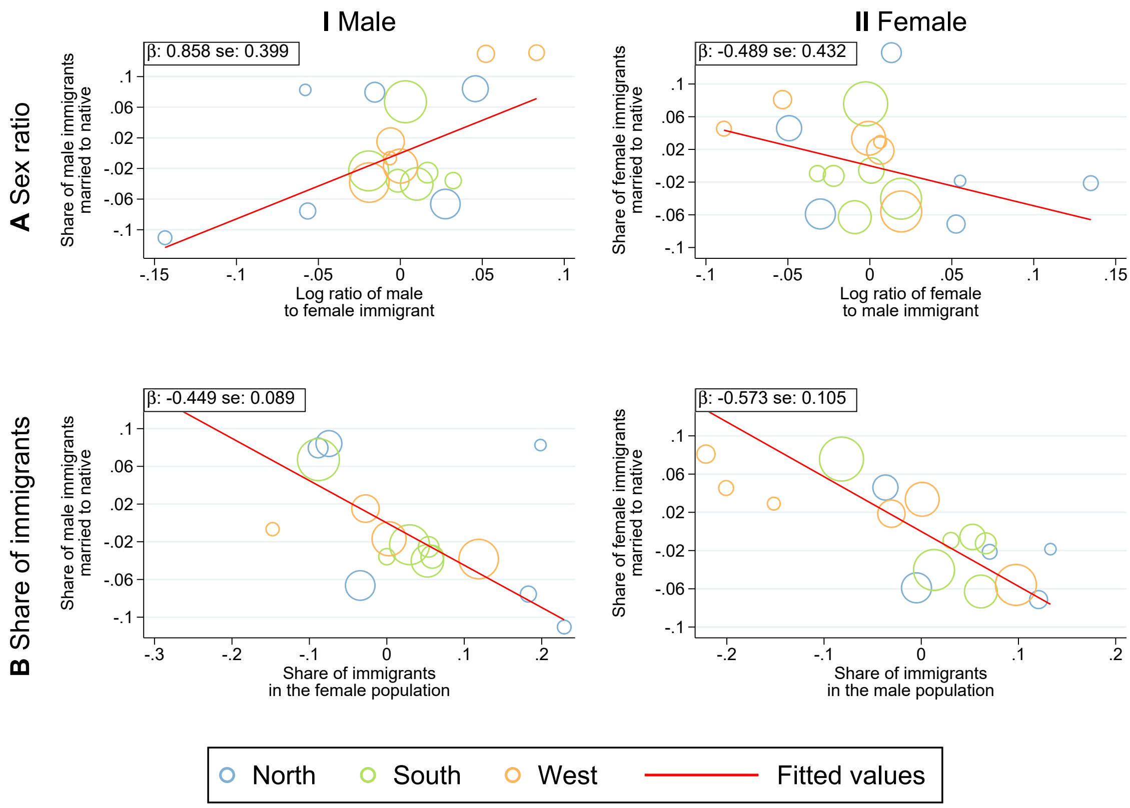
Notes: Each circle on the graph represents one group characterized by birth cohort, gender, education, and region of residence. The size of the circle corresponds to the size of the observation unit. Both variables are net of gender, education, cohort, year, age, and region fixed effects. Observations with a mean income difference above the 90th percentile and below the 10th percentile are dropped. *Source*: GSOEP 1984-2017 & German Microcensus 2006, 2010, 2015.

other countries. The diversity of immigrants' origins creates a suitable environment for analysis that allows answering the research questions of this paper.

Figure 1.1 compares region choices and the difference in mean wages between immigrants and natives. To obtain the proper comparison, I use the difference in immigrant and native probabilities of settling down in one of three German regions. By that, I measure the relative overrepresentation of immigrants in the local population. The intuition behind this exercise is as follows: immigrants from different groups settle down more often compered to natives in regions where their wages are relatively higher. The figure suggests a positive correlation. The fitted regression indicates that closing the wage gap between immigrants and natives in a particular group by 1 pp (percentage point) leads to a 0.196 (s.e. 0.090) pp increase in differences in the probability of region choice in this group. In summary, labor market conditions could be an essential factor driving immigrants' location choices.

Beyond the difference in labor market conditions, regions also differ in the local social structure. Those differences may play a vital role in determining marriage patterns. Figure 1.2 presents the correlation between the share of intermarried immigrants and two characteristics of the local marriage market: sex ratio in the immigration population (panel A); share of immigrants within the local opposite-sex population (panel B). The figure suggests a positive correlation between the intermarriage rate and the sex ratio. 1% increase in the sex ratio leads to a 0.858 p.p. (s.e. 0.399) increase in the intermarriage rate among males. The outcome indicates that a higher sex ratio leads to tougher com-

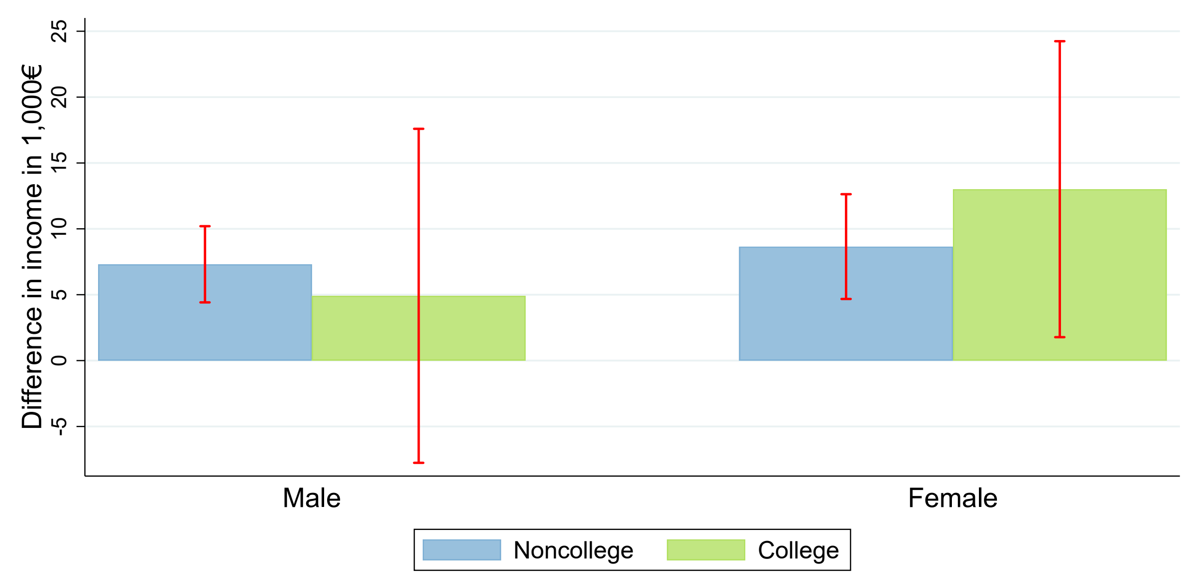
Figure 1.2: The correlation of intermarriage rate with the share of immigrants and sex ratio



Notes: Each circle on the graph represents one group. The groups are defined by birth cohort, gender, education, and region of residence. The size of the circle corresponds to the size of the group. I calculate the intermarriage rate as a share of immigrants married to natives in the group. The sex ratio is the number of males (females) per female (male). Intermarriage rate, sex ratio, and share of immigrants are net of the cohort, education, and region fixed effects. The lines represent the fitted regression lines, which slopes and their standard errors are included in the upper-left corner of each subplot. *Source*: Microcensus 2006, 2010 and 2015.

petition in the marriage market for male immigrants. As a result, it provides incentives to search for a partner outside their origin group. This conclusion does not apply to female immigrants. In their case, the correlation is negative but insignificant. It means, that competition seems to play a more important role only for male immigrants. The downer panels of the Figure 1.2 suggest that the share of immigrants in the different sex local population negatively correlates with the probability of intermarriage. Suppose a share of females (males) increases by one p.p. In that case, the percentage of intermarried male immigrants decreases on average by 0.449 p.p. (0.573 p.p) with 0.089 (0.105) s.e. It implies that the bigger pool of immigrant partners lowers the probability that an immigrant finds a partner among natives. As a result, immigrants could consider those differences between regions while deciding on their future living place.

Figure 1.3: Difference in mean income between mixed and all-immigrant households



Notes: Each bar presents the difference in means of yearly household income expressed in thousands of EUR between households where only one member is native and households where both members are immigrants. The immigrant position defines gender and education. Household income is a sum of labor and nonlabor incomes. Both incomes are net of age profile, region, and year fixed effects. The red lines represent 95% confidential intervals for the calculated mean. The standard error of the mean is calculated using clusters at the household level. *Source*: GSOEP 1984-2017.

Reports on immigrants (i.e., OECD (2020)) suggest that they differ from natives regarding labor market outcomes. Those differences manifest later in disparities in the disposable income of households. Figure 1.3 presents differences in mean income between two types of households: mixed households, cohabitation of an immigrant and native, and all-immigrant households, where both partners are foreign-born. I conduct the analysis from the point of view of an immigrant and separately by gender and education level. The figure suggests that, on average, mixed households are characterized by higher income than all-immigrant households. It links the marriage decision with economic well-being.

The correlations presented in Figure 1.1, 1.2, 1.3 are suggestive of the link between location decision, choice of partner, and labor market outcomes. Impact on the latter might influence the first two and change the final effect of the integration policy. As a result, endogenizing location and marriage choices in the equilibrium framework shed new light on the unintended effect of the pro-integration labor market policy.

1.3. Model

Why do labor market conditions influence immigrants' marriage choices? The answer to this question can be briefly described. People marry for both economic and noneconomic reasons. Regarding pecuniary motives, couples can collect more resources than single agents. The size of a family's income depends on, among others, the partners' origins. Natives tend to earn, on average, more than immigrants, so households with them have higher disposable income. Further, married agents tend to perform better in the labor market, which is associated with the marriage premium described in the literature. The premium can differ by partner's origin. The size of the additional premium received by immigrants thanks to marriage with natives might depend on the labor market integration.

Regarding noneconomic reasons for marriage, people have a taste for similarity. The taste may play an essential role in immigrants' case since they can have preferences to marry somebody who shares similar values, language, or religion. As a result, immigrants can trade economic perspectives for cultural similarity. However, the possibility of trade-off depends on the local marriage market structure. The fewer immigrants in a different sex marriageable group, the harder to marry another immigrant and comparatively easier to marry a native. The structure of the marriage market is not exogenous but depends on immigrants' location choices. While choosing where to live, immigrants take into account two aspects. First, the economic situation in the region, in particular, the level of wages. Second, the number of desirable potential partners. Changes in the labor market situation in regions can lead to a stronger or weaker spatial concentration of immigrants, directly impacting marriage market conditions.

Four things are necessary to capture the abovementioned mechanisms: (1) the model of marriage and household behavior, (2) heterogeneity in origin among agents, (3) an endogenous location decision, and (4) wages varying by agent's origin, marital status, and spouse's origin. This list motivates the following setup.

1.3.1. Set-up of the model

In the model, agents belong to a cohort of women $\mathcal{F}$ or men $\mathcal{M}$. Each agent's life is divided into three stages, indexed 1-3. At the beginning of stage 1, agent of gender $g \in \{M, F\}$ posses a human capital H. It comprises two elements: origin and education. I denote agents' origin by $o \in \mathcal{O} \equiv \{n, i\}$, where n stands for native and i stands for immigrant. Agent is also characterized by education level denoted by $e \in \mathcal{E} \equiv \{e^1, e^2\}$. As a result, human capital can be expressed as a two-element set $H \equiv \{o, e\}$. The distribution of human capital has finite support $\mathcal{H}$ of cardinality 2×2.

At stage 1, all agents first draw a vector of location preferences. Then, they make lifetime decisions regarding a region of residence. Agent chooses location, denoted by r selecting from the set $\mathcal{R} \equiv \{r^0, r^1, r^2\}$. Region choice depends on local amenities and future marriage and economic perspectives. As a result, at the end of stage 1 agent lives in the region r, where next enters a marriage market to search for a partner.

At stage 2, agents draw a vector of marital preferences and then participate in the local marriage market chosen at stage 1. The agents match based on the level of human capital (education and origin) and marital preferences in the frictionless framework. An individual can marry a person of different sex, with origin $o*$ and education $e*$. Partner's human capital is then consistently denoted by H^*. The couple can be of 16 marriage types (four types of men and four of women). I denote the married couple's type by (H, H^*), where H is the human capital of the husband and H^* is the wife's. A single household's type of man and women with human capital H is denoted by $(H, \varnothing)$ and $(\varnothing, H)$, respectively. Marriage is a lifetime decision, so the outcome of stage 2 remains forever - there is no possibility of divorce or separation.

At stage 3, agents realize their productivity and leisure shocks and observe their wages and leisure preferences. Then, all households choose the optimal consumption of private and public goods and labor supply. I assume married couples make a Pareto efficient decision.

Agent's utility splits into three parts corresponding to the model's three stages. The first part comprises the working-life utility at stage 3, derived from the consumption of goods and leisure. The second part is the utility derived from participating in the marriage market. Finally, agents derive utility from regional amenities. My description of the model is as follows. First, I define the household maximization problem at stage 3. Then I describe the marriage market, taking working-life utility as given. Finally, I provide a brief description of the location choice.

1.3.2. Working-life utility of agents

At stage 3, agents choose the optimal consumption and labor supply levels. The choice is made based on observed wages and nonlabor income. Then, agent of gender g and human capital $H = \{o, e\}$ married to agent of origin o^* in region r earn wage given by:

$$w = W_g(H, o^*, r) \cdot \varepsilon = \exp\left\{\theta_{0g}(H) + \theta_{1g}(H, o^*) + \theta_2(H, r)\right\} \cdot \varepsilon \tag{1.1}$$

where:

$$\varepsilon | g, e \sim i.i.d.\ \mathcal{N}\left(0, \sigma^2_{\varepsilon|g,e}\right) \tag{1.2}$$

The exponential expression in Equation 1.1 represents the deterministic part of agents' wages. It consists of three components. The first component $\theta_{0g}(H)$ corresponds to the agent's human capital market value. To capture the gender wage gap in labor income, I let human capital market value differ for men and women. Equation 1.1 also allows immigrants and natives with the same education to have different human capital market values for two reasons. First, immigrants can have a different intercept than natives, which captures the effect of country-specific skills, like language (Llull, 2018). Second, natives and immigrants can differ in return to education (Borjas, 1985). The second component of Equations 1.1 $\theta_{1g}(H, o^*)$ represents market value shift relate to partner origin o^*. By that, the model allows for heterogeneity in agents' wages by spouse's origin. The wage premium associated with a partner's origin varies by gender to reflect empirical facts in the literature Meng and Gregory (2005), Meng and Meurs (2009). If agent is single, so $o^* = \varnothing$, then $\theta_{1g}(H, \varnothing)$ can be interpreted as a shift in market value due to being unmarried. The third component $\theta_2(H, r)$ is introduced to capture regional variation in earnings. It makes some regions more attractive due to better labor market conditions. Agents' wages are subject to independent and idiosyncratic productivity shock ε, conditionally on gender and education, normally distributed across agents with zero mean and variance $\sigma^2_{\varepsilon|g,e}$.

Agents at stage 3 derive utility from the consumption of goods and leisure. The model has two types of goods: a public good and private good. The working-life utility has the following form:

$$u(Q, C, L) = \ln Q + \ln\left(C + \alpha(\ell_{pt} + \ell_{nw}) + \delta\ell_{pt}\right) \tag{1.3}$$

where $L = (\ell_{ft}, \ell_{pt}, \ell_{nw})$ represents agent's leisure choice, C denotes private consumption and Q corresponds to public consumption of the household. There are three available choices of leisure: full-time employment ℓ_{ft}, part-time employment ℓ_{pt} and not working ℓ_{nw}, such that $\ell_{ft} + \ell_{pt} + \ell_{nw} = 1$. If an agent is a man, then the model limits his choice to two alternatives $\ell_{ft} + \ell_{nw} = 1$, since men outside of training and education rarely actively decide to work part-time (Beham et al., 2019).

Random variable α, which represents a preference for leisure, depends on the agent's marital status $\mathbf{1}\{H^* = \varnothing\}$, gender g and human capital H. Female agents additionally have a preference shifter, denoted by δ, in case they decide to work part-time. Preference shifter δ is a random variable whose values differ for single and married females. Both α and δ are subject to the preferences shocks ξ and υ, respectively. Those shocks are uncorrelated and conditionally on gender follow the normal distribution with zero mean and variance $\sigma^2_{\xi|g}$ and $\sigma^2_{\upsilon|g}$.

Preferences satisfy the *transferable utility* (TU) property if there exists a cardinal representation of utilities, such that for all values of prices and income, the Pareto frontier is a straight line with a slope equal to -1 (Chiappori and Gugl, 2020). One can show that is true for 1.3 (by taking the $\exp u$ cardinalization). The TU property implies that household aggregate demand does not depend on Pareto weights. It means that at stage 3, a married couple (H, H^*), conditional on labor supply, chooses their optimal consumption of public goods Q and aggregated private consumption $\overline{C}$ $(= C + C^*)$ by solving the following maximization problem:

$$\max_{\overline{C},Q} \exp u(Q,C,L) + \exp u(Q,C^*,L^*) = \max_{\overline{C},Q} Q(\overline{C} + \alpha \ell_{nw} + \alpha^*(\ell^*_{pt} + \ell^*_{nw}) + \delta^* \ell^*_{pt}) \tag{1.4}$$

with respect to the budget constrain:

$$\overline{Y}^{H,H^*}(L,L^*) \equiv y_{nl}(H,H^*) + \ell_{nw} \cdot b(w) + \ell^*_{nw} \cdot b(w^*) + w_{net}(\ell,\ell^*,w,w^*) = \overline{C} + pQ \tag{1.5}$$

Household obtains income from work (w, w^*) or unemployment benefits $(b())$. The gross wages are mapped to net income using information about both partners' labor supply and income to mimic a German tax system (details in Appendix A.5.1). Buettner et al. (2019) provides evidence that households adjust their labor market choices to minimize taxation burden, which makes income mapping an important part of the model. Unemployment benefit $b()$ is defined as a function of wages to mimic the German unemployment benefit system (details in Appendix A.5.2). Households also obtain a non-labor income (conditional on both partners' human capital), denoted by y_{nl}^{H,H^*}. Household spends the budget on private consumption $\overline{C}$ and public consumption Q. The latter one they buy on the market at a price p.

Conditional on labor supply, the solutions (details in Appendix A.3.1) for public and private consumption are:

$$pQ(L,L^*) = (\overline{Y}^{H,H^*}(L,L^*) + \alpha \ell_{nw} + \alpha^* \ell^*_{nw} + \delta^* \ell^*_{nw})/2 \tag{1.6}$$

$$\begin{aligned} \overline{C}(L,L^*) &= \overline{Y}^{H,H^*}(L,L^*) - \alpha \ell_{nw} - \alpha^* \ell^*_{nw} - \delta^* \ell^*_{nw})/2 \\ &= pQ(\ell,\ell^*) - \alpha \ell_{nw} - \alpha^* \ell^*_{nw} - \delta^* \ell^*_{nw}. \end{aligned} \tag{1.7}$$

Plugging Equations 1.6 and 1.7 into the maximization problem given by Equation 1.4, provides the expression for the optimal choices of labor supply. The final maximization problem is a discrete choice problem. Each couple (H, H^*) has 3×2 choices of labor supply, formally:

$$\max_{L,L^*} pQ^2(L,L^*) \tag{1.8}$$

The single maximization problem at Stage 3 follows the one presented for couples. Appendix A.3.2 explains the single maximization problem and describes its solution.

At stage 2, agents do not know the future realization of the productivity and leisure preference shocks. Define $C^* = (\overline{Y}^{H,H^*}(L,L^*) - \alpha \ell_{nw} - \alpha^*(\ell^*_{nw} + \ell^*_{pt}) - \delta^* \ell^*_{pt})/2 - C$, then the ex ante efficient allocation is given by:

$$\max_{C} \mathrm{E}\, u + \mu \,\mathrm{E}\, u^* \tag{1.9}$$

The solution to this problem is a set of Pareto efficient allocations given by:

$$\exp\{\mathrm{E}\, u\} + \quad \exp\{\mathrm{E}\, u^*\} = \frac{1}{1+\mu}\exp\{\Psi(H,H^*,r)\} + \frac{\mu}{1+\mu}\exp\{\Psi(H,H^*,r)\} \tag{1.10}$$

$$U_g(H,H^*,r) + \quad U_{g^*}(H^*,H,r) = \exp\{\Psi(H,H^*,r)\} \equiv \overline{U}(H,H^*,r) \tag{1.11}$$

where:

$$\Psi(H,H^*,r) \equiv \ln p + \int \ln Q^2(H,H^*,r,\boldsymbol{\varepsilon},\boldsymbol{\upsilon},\boldsymbol{\zeta})dF(\boldsymbol{\varepsilon},\boldsymbol{\upsilon},\boldsymbol{\zeta}) \tag{1.12}$$

U at stage 3 represents the agent's expected working-life utility from the union (H,H^*) generated at stage 3. Similarly, the function $\overline{U}(H,H^*,r)$ represents the total economic value generated by the couple (H,H^*). It is worth stressing that $\overline{U}(H,H^*,r)$ is the function only of the partners' human capital and region of residence. The detail derivation of both functions are in Appendix A.3.1).

For single agents, the ex-ante (again, before the realization of the productivity and leisure preference shocks) Pareto efficient set of allocation is defined as:

$$U_g(H,\varnothing,r) = \exp\{\mathrm{E}\, u\} \tag{1.13}$$

Note that $U_g(H,\varnothing,r)$ refers to the same cardinalisation as in Equation 1.9.

1.3.3. Marriage market

At stage 2, agents enter the local marriage markets. They decide whom to marry or to stay single based on preferences and expected utility at stage 3. Let a set of male (female) with human capital H (H^*) living in region r be $N_M^{H,r}$ ($N_F^{H^*,r}$). To identify parameters in the marriage market, I follow the separability assumption in Galichon and Salanié (2021). It states that the total value generated by marriage is a sum of two elements: systematic and idiosyncratic components.

The systematic component consists of an expected economic value obtained by marriage at stage 3 (given by the Equation 1.9) and taste for similarity (or rather distaste for dissimilarity). The letter one captures the distaste for the divergence in origin (denoted by $\phi_1|o^*-o|$) and the distaste for the difference in education (denoted by $\phi_2|e^*-e|$). Agents (conditional on their human capital and gender) also have a taste for being single.

An idiosyncratic component is the second element of the marriage surplus. Let $\boldsymbol{\omega} = (\omega_{H^*,r} : H^* \in \mathcal{H} \cup \{\varnothing\})$ denote the payoff vector of individual, which represents subjective satisfaction in region r from being married to a person with human capital H^* or staying single. The second part of the separability assumptions stands that an individual draws vector $\boldsymbol{\omega}$ from the probability distribution $\boldsymbol{Q}_g^H$ conditional on gender. It additionally assumes that $\max_{H^* \in \mathcal{H} \cup \{\varnothing\}} |\omega_{H^*,r}|$ have finite expectations under $\boldsymbol{Q}_g^H$.

Formally, the total gain generated by the match between a man with H and a woman H^* living in the region r is:

$$\overline{\Gamma}(H, H^*, r) = \Gamma_M(H, H^*, r) + \Gamma_F(H^*, H, r) \tag{1.14}$$

where $\Gamma_M(H, H^*, r)$ and $\Gamma_F(H^*, H, r)$ are partners' individual utilities.

Agents find their preferred partners by maximizing utility. The preferences are characterized by the transferable utility, which means that the surplus given by the Equation 1.14 is fully divided between spouses. The Pareto weight μ associated with the initial log cardinalization drives the division of the future expected working-life utility. Formally:

$$\Gamma_M(H, H^*, r) = \overline{U}(H, H^*, r) - \tau(H, H^*, r) + \phi_1|o^* - o| + \phi_2|e^* - e| \tag{1.15}$$

$$\Gamma_F(H^*, H, r) = \tau(H, H^*, r) + \phi_1|o - o^*| + \phi_2|e - e^*| \tag{1.16}$$

where:

$$\tau(H, H^*, r) = \frac{\mu(H, H^*, r)}{1 + \mu(H, H^*, r)} \overline{U}(H, H^*, r), \quad \mu(H, H^*, r) > 0$$

Pareto weights act as a price that ensures market clearing. This assumption, together with the fact that idiosyncratic shocks are assumed to be independent across two partners, allows me to identify the transfers between agents in the marriage market using marriage outcomes (see Proposition 1 in Galichon and Salanié (2021)).

In the marriage market, some agents match while others do not. A single agent of gender g with human capital H derives utility of the following form:

$$\Gamma_g(\varnothing, r) = U_g(H, \varnothing, r) + \phi_{0H} + \omega_{\varnothing,r} \tag{1.17}$$

The mapping of who marries whom and who stays single is a match. In the model, I consider the stable match - a match under which no agents have an incentive to deviate from the equilibrium. Formally, the stable match is defined as follows:

Definition 1. A *stable matching* for a marriage market in region r is a triple $(N_M^r, N_F^r, \boldsymbol{\Gamma}(r))$, where N_M^r (N_F^r) is a set of men (women) living in region r and $\boldsymbol{\Gamma}(r)$ is a set of payoffs for any men and women, such that for any $H, H^* \in \mathcal{H}$ in r:

1. $\Gamma_M(H, H^*, r) \geq \Gamma_M(H, \varnothing, r)$ for all men

2. $\Gamma_F(H^*, H, r) \geq \Gamma_F(H^*, \varnothing, r)$ for all women

3. $\Gamma_M(H, H^*, r) + \Gamma_F(H^*, H, r) \geq \overline{\Gamma}(H, H^*, r)$ for all men and women

4. $\Gamma_M(H, H^*, r) + \Gamma_F(H^*, H, r) = \overline{\Gamma}(H, H^*, r)$ for all matched couples

The first two conditions refer to the *individual rationality* assumption - none of the matched agents can be worse off than while staying single. Condition number 3 refers to the idea of *blocking pairs*. A matching is stable if there are no two agents of the opposite sex such that while matching, they are better off than in their current matching. The last condition states that the sum of individual utilities from marriage equals the total value generated in the match. It is a direct consequence of *transferable utility* assumption.

In theory, Pareto weight μ can be match specific. However, following Chiappori et al. (2018), one can show that μ is specific for a combination partners' human capital (H, H^*). Formally:

Proposition 1. *In a stable match, consider two couples (H, H^*) and (H', H'^*) living in the same region r, such that $H = H'$ and $H^* = H'^*$. Then the Pareto weight is the same for both couples.*
Proof.
From the condition 4 (no blocking pairs) and 5 (transferable utility) of Definition 1 and Equation 1.14, we have:

$$
\begin{aligned}
\Gamma_M(H, H^*, r) + \Gamma_F(H^*, H, r) &= \overline{U}(H, H^*, r) + 2 \cdot (\phi_1|e^* - e| + \phi_2|o^* - o|) + \omega_{H^*,r} + \omega^*_{H,r} \\
\Gamma_M(H, H^*, r) + \Gamma_F(H'^*, H', r) &\geq \overline{U}(H, H'^*, r) + 2 \cdot (\phi_1|e'^* - e| + \phi_2|o'^* - o|) + \omega_{H'^*,r} + \omega'^*_{H,r} \\
\Gamma_M(H', H'^*, r) + \Gamma_F(H'^*, H', r) &= \overline{U}(H', H'^*, r) + 2 \cdot (\phi_1|e'^* - e'| + \phi_2|o'^* - o'|) + \omega'_{H'^*,r} + \omega'^*_{H',r} \\
\Gamma_M(H', H'^*, r) + \Gamma_F(H^*, H, r) &\geq \overline{U}(H', H^*, r) + 2 \cdot (\phi_1|e^* - e'| + \phi_2|o^* - o'|) + \omega'_{H^*,r} + \omega^*_{H',r}
\end{aligned}
$$

Then subtracting the first two and the last two equations gives:

$$\omega^*_{H,r} - \omega'^*_{H,r} \geq \Gamma_F(H^*, H, r) - \Gamma_F(H'^*, H', r) \geq \omega^*_{H',r} - \omega'^*_{H',r}$$

which leads to the conclusion:

$$\Gamma_F(H^*, H, r) - \omega^*_{H,r} = \Gamma_F(H'^*, H', r) - \omega'^*_{H,r}$$

It means that the difference between the utility obtained by the wife and her idiosyncratic component is constant across agents with the same human capital. As a result, μ depends only on partners' human capital (H, H^*). □

Proposition 1 shows that, in a stable matching, an individual's utility is simply a sum of three elements: (1) an idiosyncratic shock, (2) noneconomic preferences for partner human capital, (3) endogenously determinate on the labor market share of future economic gain generated by the household. Thanks to that, it is possible to express each individual's problem as a discrete choice problem. It is described in the following proposition:

Proposition 2. *In a stable match, a utility of a man with H satisfies:*

$$\tilde{\Gamma}_M(H, r) = \max_{H^* \in \mathcal{H} \cup \varnothing} \Gamma_M(H, H^*, r) + \omega_{H^*, r} \tag{1.18}$$

and utility of female j satisfies:

$$\tilde{\Gamma}_F(H^*, r) = \max_{H \in \mathcal{H} \cup \varnothing} \Gamma_F(H^*, H, r) + \omega^*_{H, r} \tag{1.19}$$

The Proposition 2 states that the discrete choice problem is determined by utility transfers between agents μ, preferences for the difference in partners' education and origin, and individual idiosyncratic shock. μ is exogenous from the agent's perspective. It acts as a price on the marriage market and ensures that no one has an incentive to deviate from stable matching.

Let's assume that agent draws $\boldsymbol{\omega}$ from Extreme Value Type I distribution with variance $\sigma_\omega^{g,H}$. Then, the probability that an agent with H living in the region r marries an agent with human capital H^* is:

$$\mathrm{P}_r\left(H^* | H, r\right) = \frac{\exp\left\{\Gamma_M(H, H^*, r)/\sigma_\omega^{g,H}\right\}}{\sum\limits_{H^* \in \mathcal{H} \cup \varnothing} \exp\left\{\Gamma_M(H, H^*, r)/\sigma_\omega^{g,H}\right\}} \tag{1.20}$$

At the beginning of Stage 2, agents do not know their idiosyncratic preferences. Using the distribution of $\boldsymbol{\omega}$, the expected utility from stage 2 is given by:

$$\hat{\Gamma}(H, r) = \mathrm{E}\left[\tilde{\Gamma}_M(H, r)\right] = \ln\left(\sum_{H^* \in \mathcal{H} \cup \varnothing} \exp\{\Gamma_M(H, H^*, r)/\sigma_\omega^H\}\right)^{\sigma_\omega^H} + \gamma \tag{1.21}$$

where γ is an Euler constant.

1.3.4. Location choice

At Stage 1, agents decide about their future location. There are three possible location choices. Each region is associated with a level of regional amenities, which capture the region's (unrelated to marriage and labor market) attractiveness, e.g., environmental conditions, crime level, transportation system, or general economic situation.

Formally, agents choose their region of residence as follows:

$$r = \arg\max_{r \in \mathcal{R}} \hat{\Gamma}(H, r) + \beta \times Z_r + \eta_r \tag{1.22}$$

where Z_r represents a vector of regional amenities. $\hat{\Gamma}(H, r)$ is defined like in Equation 1.21. The choice of the region takes into account both the returns in the labor market and the marriage market structure. Individuals at that point do not know their idiosyncratic components. This assumption corresponds to the situation where agents are unaware of their marital preferences and rather learn about them while meeting new people and dating. η is an idiosyncratic shock, which measures the subjective preferences of agents towards a given region.

If η's are Extreme Type I value distributed with variance σ^g_η, then the probability that agent with H settles down in region r is given by:

$$P(r|H) = \frac{\exp\{(\hat{\Gamma}(H,r) + \beta \times Z_r)/\sigma^g_\eta\}}{\sum_{r' \in \mathfrak{R}} \exp\{(\hat{\Gamma}(H,r') + \beta \times Z_{r'})/\sigma^g_\eta\}} \tag{1.23}$$

Finally, the structure of that stage is a sequential game: agents choose first where they would like to live, but their future utility depends on the distribution of human capital on both sides of the marriage market in the chosen location.

1.4. Data and Estimation

1.4.1. Data

This subsection briefly discusses the data and sample used in the estimation. More detailed descriptive statistics of the sample are in Appendix A.1. There are two primary sources of the data used in this paper. Wages and labor supply choices are estimated using the German Socio-Economic Panel (GSEOP) data for 1984 - 2018. The primary sample includes all males and females aged between 25 and 55[8]. Those enrolled in school or who changed their region of residence are excluded. The final sample contains only singles observed past age 30 to avoid underestimating the marriage rate. For married couples, I only include observations from the first marriage.

Regarding the subsample of immigrants, I exclude those who married before migration since they do not participate in the marriage market in Germany. Additionally, I exclude observations from East Germany due to two reasons. First, the share of immigrants in East Germany is very low, close to zero. Second, economic and law conditions differ in West and East Germany, which can impact household choices. The final dataset contains information on: education (college vs. noncollege), origin (native or immigrant), labor and nonlabor income, and employment for 94,003 households, of which 78,787 are couples, 6,595 are male singles, and the remaining 8,621 are single women.

The second data source is Microcensus for 2006, 2010, and 2015. The survey collected data with a sampling fraction of 1% of the persons and households in Germany. Due to the size and representativeness, the sample constructed from Microcensus data is used to calculate marriage and region choice probabilities. I divided the sample into 10-year

[8]For couples, the female age and year of birth are reference one.

birth cohorts: agents born in the '50s, '60s, and '70s. Then the marriage market and location choice are estimated separately for each cohort.

In the empirical analysis, labor supply decisions are classified into three groups. Full-time workers are those agents who report working at least 35 hours per week. Part-time workers are all female workers who work from 1 to 35 hours per week. All remaining agents are assigned as not-working. The reported wages below the 1st or above the 99th percentile are trimmed to limit the impact of the extreme observation on the estimation result. The model is static, so wages are net of time and age effects. Nonlabor income is replaced, but its estimates net of year and age effects are based on household human capital. Two levels of education correspond to college and noncollege graduates. All agents who were born outside of Germany are qualified as immigrants.

1.4.2. Outline of the estimation

In this subsection, I discuss the three-step procedure to estimate the model: (1) I estimate outside the model age-year profile and nonlabor income, (2) the estimation proceeds with parameters associated with wages and labor supply choices, (3) the marriage market and location choice parameters are estimated.

Outside the model estimation

I start with an estimation of the nonlabor income of singles and couples on their human capital and region of residence. Predicted nonlabor income is used in the budget constraint at Stage 3. Due to the static form of the model, wages used at Stage 3 are net of age and year effects. It requires estimation of age profile with year fixed effects. I use a control function approach as in Heckman (1979) to allow for endogenous selection to employment. Residuals from the nonlabor income regression and the number of children younger than five are used as exclusion restrictions. To clear wages from age and year effects, I estimate the following equation:

$$\ln \tilde{w} = \ln w + \gamma_0^H age + \gamma_1^H age^2 + \gamma_2^H age^3 + \gamma_4^{year} + \lambda(z^w \beta^w) + \varphi \tag{1.24}$$

where $\ln w$ is a wage given by Equation 1.1, γ_4^{year} is a year dummy, and $\lambda(z^w \beta^w)$ is a control function for employment. Vector z^w includes individual residuals from the nonlabor income regression, number of children younger than five, age polynomial, and year dummies. Having estimated the age profile with year dummies, I can replace reported wages with predicted ones. I predict the wages for the age 45 (average age in the sample), keeping their level as in 2005 (sample median). At stage 3, net wages are taken as given. The procedure to estimate wage parameters and leisure preferences is provided below.

Wages and leisure preferences at Stage 3

The utility of a household (H, H^*) at stage 3 is given by the Equation 2. Individuals' preferences for leisure α depend on their human capital and marital status. They are also subject to the leisure preferences shock, formally:

$$\alpha = \alpha_{0g}^{H} + \alpha_{1g}^{H} \cdot \mathbf{1}\{H^* \neq \varnothing\} + \zeta \tag{1.25}$$

Additionally, females who decide to work part-time have a preference shifter δ. It is a random variable that measures how much female preferences for leisure are different if they decide to work part-time. δ depends on agents' marital status and is subject to the normally distributed shock conditional υ, formally:

$$\delta = \delta_0 + \delta_1 \cdot \mathbf{1}\{H^* = \varnothing\} + \upsilon \tag{1.26}$$

Wages in the model are as described by Equation 1.1 and are subject to the productivity shock ε. Productivity and preferences for leisure shocks are drawn from the corresponding distributions at the beginning of Stage 3 (so after the marriage market). It allows me to treat observed marital patterns as given and estimate the postmarital part of the model separately. All wage and leisure preference parameters are estimated using the method of simulated moments. I use 192 moments, which include (1) means, variances, and quantiles of the wage distribution and probability of working and part-time, all by gender, marriage status, and human capital, (2) means and variances of wage distribution by gender, own and partner's human capital to identify the impact of marriages on agents' earnings, (3) means of wage by region, gender and human capital level to identify regional differences in wages. Appendix A.6 contains the complete list of data and simulated moments.

Marriage market and location choices

This subsection briefly describes the estimation procedure for the parameters of the marriage market and location choices. First, I present an estimation strategy for marriage market parameters. Then, I describe the construction of the region amenity index. Finally, I briefly discuss the estimation of location choice parameters.

The marriage market is estimated following the approach by Choo and Siow (2006). In the model, there are two levels of education (college and noncollege) and two origins (native and immigrant). As a result, each agent chooses a partner among four alternatives. In the marriage market, preferences for partners are observable for all participants. Wage shocks and leisure preferences are unknown, so agents match based on expectations. I have estimated in earlier steps the wages and leisure preferences parameters. It allows me to compute the expected economic component for all possible matches defined by the human capital and for singles of all types in each region. I use these estimates to identify the Pareto weights in this step. The estimation procedure treats Pareto weight and marital preferences parameters as unknown. Marital preferences are identified based on observed choices. Variances of the marital shock are identified based on variations in the expected economic component across regions.

I estimate the model separately for each region and cohort. First, I derive a set of quasi-demand and quasi-supply functions following Choo and Siow (2006) (see Appendix A.4 for details). Since the value of $\overline{U}(H, H^*, r)$ is calculated using estimates from stage 3, the set of Pareto weights is fully identified using the quasi-demand functions (male choices). However, I can also use quasi-supply equations (female choices), which leads to overidentification and allows me to identify the remaining marriage market parameters. I use a minimum distance estimator. The algorithm searches for parameters that minimize the distance between observed quasi-demand and quasi-supply functions and ones implied by the model.

The location choice in the model depends on the region amenities index. The amenity index should ideally capture the whole bundle of amenities, accurately measuring the quality of living in the region. Region amenities index is calculated following the procedure provided by Diamond (2016). First, I collect data on eight different amenities in 16 German regions. The data captures the period from 1971 to 2000, corresponding to the time when cohorts from my sample were aged 21-30. It is a period of life when people are most likely to make their lifetime decision regarding their place of living and marriage. Then, I divide those amenities into four groups: transportation, environmental, crime, and economy. Next, I create amenity subindices using principal component analysis (PCA). Then I use those subindices to calculate an overall amenity index. The final amenity index is aggregated into three regions used in the model using population as a weight.

Table 1.1 presents the loadings on each amenity subindex and the final overall amenity index. The transportation index negatively weighs the number of passengers but positively length of highways per km^2. It suggests that a single measure of transportation can approximate the development of road networks, which leads to a decrease in the use of public transport. The environmental subindex positively loads the share of forest in the region and the number of national parks. The crime index puts a positive weight on the number of crimes and the number of severe crimes. Finally, the economy index positively weighs GDP per capita and employment, indicating that higher GDP per capita is associated with more jobs.

To create an overall amenity index, I combine all described above subindices. The index accurately places a positive loading on transport, the environment, and the economy. On the other hand, the index weighs negatively on crimes. Intuitively, the amenity index comoves with the safety level. To sum up, a single amenity index constructed based on several subindices represents their common component well.

When choosing the region of residence, agents consider future marriage and labor perspectives. I use estimation parameters from stages 2 and 3 to calculate the continu-

Table 1.1: Principle Component Analysis for amenity indices

	Loading	Unexplained variance
Transportation subindex		
Number of passengers in public transport per capita	-0.707	0.220
Length of highways per km^2	0.707	0.220
Environmental subindex		
Forest area in %	0.707	0.162
Number of national parks	0.707	0.162
Crime subindex		
Number of crime cases per capita	0.707	0.282
Number of sever crime cases per capita	0.707	0.282
Economy subindex		
GDP per capita	0.707	0.120
Employment per capita	0.707	0.120
Overall amenity index		
Transport	0.468	0.427
Environmental	0.574	0.138
Crime	-0.550	0.208
Economy	0.386	0.610

Notes: All amenity data measured in standard deviations for the cohort. See Appendix A.2 for detailed description of amenity data.

ation value of living in a region for each human capital and cohort. Together with the calculated amenity index, it allows me to estimate parameters associated with location choices. I use the method of moments estimator by minimizing the distance between observed region choice probabilities and the one implied by the model.

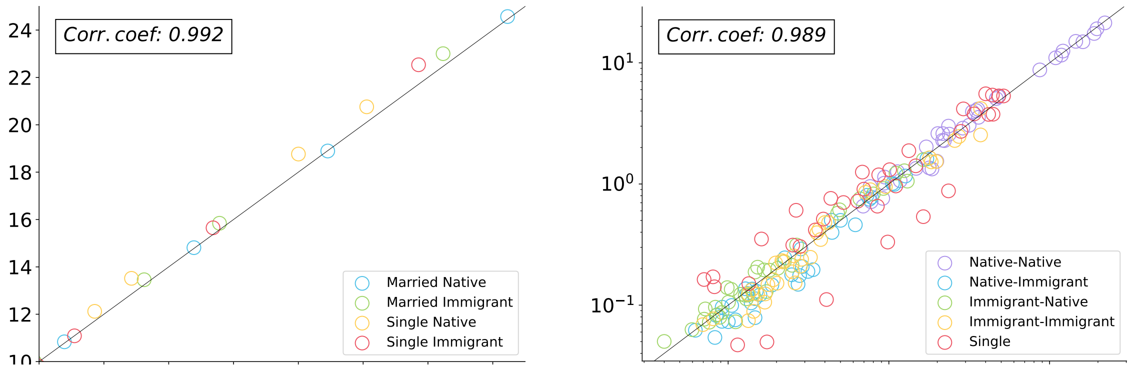
1.4.3. Estimation outcomes

This Section first discuss model's fit. Then, I show estimation of Stage 3, which are those in Equation 1.1, 1.25 and 1.26. Passing then onto a presentation of estimated parameters of Stage 1 and 2 included in 1.20 and 1.23.

Model fit

The model's fit regarding wages is presented in Figure 1.4a. The correlation between observed and predicted wages is equal to 0.992. It means that the model quite well explains

Figure 1.4: Fit of the model



(a) for wages (b) for marriage and location

Notes: Panel A of Figure 1.4 presents the correlation between observed and predicted wages. Each circle corresponds to a different group defined by marital status, gender, education, and origin. The calculated correlation coefficient is equal to 0.992. Panel B of Figure 1.4 presents the correlation between observed and predicted marriage rates. The marriage rate is defined as a share of individuals by human capital being married to a type of partner (or staying single) in the total cohort population. The cohorts are agents born in the '50s, '60s, and '70s. The marriage rates are multiplied by 100.

variation in agents' wages by observable characteristics. The prediction of the model regarding the marriage and location choices are displayed in Figure 1.4b. The model tends to under- and over-predict the share of single agents, although the general marriage patterns are captured relatively well. The correlation coefficient between observed and predicted marriage rates is equal to 0.989. All targeted moments and their fits for all estimation steps are included in Appendix A.6.

Wage equations and leisure preferences

Table 1.2 presents estimated parameters in the wage equation associated with human capital and marriage. In line with the literature findings, there is a positive return to education for all groups of agents. The wages of native agents are higher than immigrants with the same education level. It can suggest that human capital is not fully transferable between countries. When it comes to parameters associated with marriage, they are positive for married men (except for college-educated immigrants) and negative for women (except for noncollege immigrants). It would suggest that women are penalized in the labor market when married. In some cases, the value of the marriage premium depends on the partner's origin. However, the effect is not strong (coefficients are insignificant at a confidence level of 0.05), which could suggest that the partner's origin impacts total household labor income but does not influence the agent's wage.

These estimates allow the model to predict some regularities in wages, migration, and marriages established in the literature. Returns to education are quantitatively similar to those presented in Card (1999). Psacharopoulos and Patrinos (2004) finds that, on

Table 1.2: Wage parameters - human capital and marriage premium

	Native noncollege	Native college	Immigrant noncollege	Immigrant college
Male				
Constant	2.378	2.817	2.140	2.653
	(0.017)	(0.026)	(0.054)	(0.089)
Married	0.178	0.210	0.214	-0.000
	(0.012)	(0.022)	(0.043)	(0.094)
Married to immigrant	0.006	0.050	0.006	-0.012
	(0.019)	(0.037)	(0.028)	(0.080)
Female				
Constant	2.319	2.813	1.987	2.436
	(0.026)	(0.021)	(0.060)	(0.089)
Married	-0.146	-0.057	0.054	-0.071
	(0.013)	(0.019)	(0.043)	(0.076)
Married to immigrant	0.017	0.029	-0.034	0.058
	(0.027)	(0.056)	(0.033)	(0.085)

Notes: Standard errors in parentheses. Wages are expressed in log values, deflated using 2005 prices. Base category: single living in North.

average, females have higher education returns than males. The estimated parameters of the model are in line with this finding when it comes to the married population. I find lower returns to education among immigrants confirm the less-than-perfect international transferability of human capital found in, i.e., Chiswick and Miller (2009). The effect of intermarriage on wages is positive but insignificant, which supports the hypothesis by, i.e., Kantarevic (2004) and is contrary to the finding of Meng and Gregory (2005).

Estimation outcomes of wage equations have significant implications for the marriage market. The estimated parameters suggest that when immigrants decide to marry a native compared to immigrants, they can count on higher household income in the future. On the other hand, marriage with an immigrant does not pay off for natives - mixed household labor income could be lower than in the case of all-native households. The final size of the economic gain of immigrants from a marriage with a native depends on their bargaining power within the household. A higher Pareto weight could compensate for the lower household income. Shifting part of the income from immigrants to natives increases the relative attractiveness of immigrants from a native perspective.

Table 1.3 contains outcomes of the region fixed effect estimation in wage equation. First, on average, agents earn the highest wage in the South region and the lowest in the North. Second, returns to human capital are not homogeneous across regions, and immigrants experience stronger variations. Noncollege immigrants earn more in the South than in the North on average by 13.4%. The effect for college-educated immigrants is even stronger and equals 28.3%. It could suggest that the South is an attractive migration destination for foreign-born individuals. Agents who live in region West also earn more than those living in region North. However, the positive effect is weaker (3.2% vs. 13.4% for noncollege immigrants) and more homogeneous than in the region South (the difference between estimated coefficients is smaller).

Table 1.3: Wage parameters - differences across regions

Region	Human capital			
	Native noncollege	Native college	Immigrant noncollege	Immigrant college
South	0.075	0.090	0.134	0.283
	(0.010)	(0.017)	(0.032)	(0.084)
West	0.046	0.075	0.032	0.138
	(0.011)	(0.017)	(0.033)	(0.083)

Notes: Standard errors in parentheses. Wages are expressed in log values, deflated using 2005 prices. Base category: North.

I present estimated parameters associated with leisure preferences in Table 1.4. Single males have, on average lower leisure preferences than married ones for all human capital types, while the opposite is true for females. This difference can be partially explained by a higher number of children among married couples in comparison to singles. Higher fertility may cause married men to take a job more often, while women drop from the labor market to take care of children.

Interestingly, on average, immigrants have higher leisure preferences than natives. It could be related to several things. First, they may produce more at home. Home production is not included in the model so it may be partially captured by the parameter α. Home production is also more important for immigrants, since they may face problems to buy ethnic products on the market. Second, immigrants in the data have, on average, more children than natives, which can also contribute to higher leisure preferences. Part-time work shifter δ is positive for females and higher for single ones. It could suggest that in this group, mixing work with leisure (potentially taking care of children) is an additional source of utility.

Results presented in Table 1.2 and 1.4 suggest significant differences between natives and immigrants regarding wages and leisure preferences exist. Those differences are demonstrated in the economic value generated by single households at Stage 3. Table 1.5 presents it for singles living in region West. Economic value is increasing in education. It is also higher for natives than for immigrants. Similar patterns are found in Table 1.6, which presents the expected marriage economic value for couples living in the region West for all 16 possible combinations of human capital. There are two important conclusions

Table 1.4: Leisure parameters - preferences for not-woring and part-time working

	Male		Female	
	constant	married	constant	married
Pref. for leisure, α				
Native noncollege	1.591	-1.859	0.097	1.690
	(0.117)	(0.167)	(0.394)	(0.626)
Native college	0.150	-1.759	-1.451	1.818
	(0.347)	(0.301)	(0.660)	(0.691)
Immigrant noncollege	2.974	-1.574	1.706	1.877
	(1.410)	(1.404)	(0.329)	(0.489)
Immigrant college	2.776	-0.589	0.801	2.893
	(0.867)	(0.936)	(0.633)	(0.850)
Part-time work shifter, δ			6.217	-5.335
			(1.799)	(1.809)

Notes: Standard errors in parentheses. Column 'married' presentes change in parameters value for married.

from this. First, from a female point of view, marriage with a native generates higher economic gain than marriage with an immigrant with the same education level. The reason is that most men work (so their leisure preferences matters less), and native men have higher wages than an immigrant. Second, from an immigrant male point of view, marriage with a native also generates higher economic gain than marriage with an immigrant with the same education level. Interestingly, it is not true for native men. The reason could be that immigrant women have lower wages and higher preferences for leisure, so when they marry a native, they more often do not work and generate more utility from this choice than native women.

Table 1.5: Economic value of staying single in region West

	Human capital			
	Native noncollege	Native college	Immigrant noncollege	Immigrant college
Male	24.81	51.79	16.38	49.06
Female	31.41	49.45	29.98	38.78

1.4.4. Marriage market and location choice

This subsection briefly describes the estimation outcomes of parameters associated with the marriage market and location choice. Table 1.7 contains estimated parameters for

Table 1.6: Economic value of marriage - West

Male human capital	Female human capital: Native noncollege	Native college	Immigrant noncollege	Immigrant college
Native noncollege	76.65	96.81	79.21	98.42
Native college	134.31	158.57	151.37	176.05
Immigrant noncollege	69.36	92.83	67.47	92.72
Immigrant college	101.11	129.73	97.35	128.09

marital preferences. The parameters are estimated separately for each birth cohort - agents born in the '50s, '60s, and '70s of the XX century. The left-hand side of the table presents estimates of preferences for being single. They are higher for college-educated agents and natives. Also, preferences for being a single increase over time, which is in line with the data that suggest that share of people who decide to stay single is increasing. The right-hand side of Table 1.7 contains estimated parameters associated with a taste for dissimilarity in education and origin. Estimates are negative for differences in education level and origin, but the distaste for dissimilarity in origin is higher. It could be correlated with higher social norms, which need to be broken when one marries a person from a different origin group. Interestingly, the interaction term is positive and offsets the negative effect of a difference in origin and education in some parts. It could suggest that breaking both norms is associated with smaller negative tastes. All these estimates suggest that the agent prefers to marry people of the same origin and education. As a result, they have a strong tendency to trade better economic outcomes for similarities. It leads to highly positive assortative mating in the marriage market.

Table 1.7: Marriage market parameters - taste for staying single and similarity

ϕ_{0H}	Native noncollege	Native college	Immigrant noncollege	Immigrant college	$\phi_1\|e-e^*\|$	$\phi_2\|o-o^*\|$	$\phi_3\|e-e^*\| \cdot\|o-o^*\|$
'50	10.220	65.184	10.622	49.522	-20.042	-28.990	3.466
	(2.313)	(4.565)	(6.629)	(4.688)	(2.224)	(2.328)	(0.793)
'60	17.575	69.070	14.387	52.878	-18.381	-28.780	4.449
	(1.941)	(4.635)	(6.438)	(4.636)	(2.127)	(2.390)	(1.041)
'70	20.851	73.082	16.356	58.155	-18.260	-25.045	2.888
	(1.795)	(4.613)	(6.349)	(4.355)	(1.996)	(2.116)	(0.671)

Notes: Asymptotic bootstrapped standard errors in parentheses. The left-hand side of the table presents estimates of the taste for staying single by cohort and human capital. The right-hand side of the table presents estimates of taste for similarity by cohort.

Agents choose a partner based on individual marital preferences and the share of the economic gain generated by the couple, which corresponds to the Pareto weight of the collective household model. The share is unique for each type of couple living in a region. It depends not only on the agent's human capital but also reflects the relative scarcity of spouses. Therefore, it depends on the entire human population distribution in the given region. The share acts as a price that clears the marriage market. Table 1.8 presents a share of the gains from a marriage that belongs to women for couples born in the '60s and living in the West region. A higher education level generally correlates with a higher share of future utility. If a college-educated immigrant woman marries a noncollege immigrant, she gets 81% of welfare. Similarly, if a college-educated native man wants to marry a college-educated native woman, her share will be higher than that of noncollege-educated native women by around 20 pp. When the patterns for education are clear, it is not valid for origin. On average, women can extract a higher share from native men than immigrants of the same education level. It is also true for men regarding native and immigrant women. It reflects the scarcity of immigrants in comparison to natives. The higher share attributed to immigrant women compared to immigrant men suggests that immigrant women have better opportunities and stronger bargaining power while intermarrying than immigrant men.

Table 1.8: Sharing rule among couples born in '60s, living in region West

Male human capital	Female human capital: Native noncollege	Native college	Immigrant noncollege	Immigrant college
Native noncollege	0.567	0.797	0.698	0.859
	(0.015)	(0.009)	(0.012)	(0.021)
Native college	0.330	0.522	0.355	0.489
	(0.009)	(0.009)	(0.023)	(0.017)
Immigrant noncollege	0.484	0.685	0.661	0.805
	(0.013)	(0.032)	(0.024)	(0.035)
Immigrant college	0.218	0.440	0.330	0.516
	(0.028)	(0.021)	(0.032)	(0.030)

Notes: Asymptotic standard errors in parentheses are computed using the bootstrap method.

Table 1.9 presents estimation of parameters associated with location choice. The parameter associated with region amenities is positive for all human capital levels. It is the strongest for noncollege immigrants. It suggests that regional amenities play the highest role for this group when they choose their future place of living. It causes agents with a higher taste for amenities to be less sensitive to changes in wages and marriage market conditions.

Table 1.9: Taste for amenities

	Native noncollege	Native college	Immigrant noncollege	Immigrant college
β	0.106	0.089	0.174	0.049
	(0.004)	(0.008)	(0.006)	(0.020)

Notes: Standard errors in parentheses.

1.5. Policy scenarios

This section presents the analysis of two policy scenarios. First, I briefly describe the scenarios. Next, I present changes to single and intermarriage rates under each scenario and discuss the consequences of those changes for the integration of immigrants. Further, I show how immigrants' concentration transforms and how it impacts integration. Finally, I present income inequality and welfare changes in each case with and without adjustment in marriage and location choices.

Policymakers might target immigrants' integration through various labor market policies. Reducing the immigrant-native wage gap is indispensable to holistic integration (Lehmer and Ludsteck, 2015). In scenario I, the government introduces a policy that directly impacts immigrants' wages. As a result, the immigrant-native wage gaps across genders and all education levels decrease by 25%. It is an equivalent to an average increase in wages of immigrant men by 4.6% and of immigrant women by 7.5%. Instead of a universal closer to the wage gap, a government might prefer to focus on regional wage variation. In scenario II, the introduced policy leads to the closing of the difference between the best region (South) and the remaining regions (North and West) by increasing the average wage in the remaining regions by 50% of the initial size of the gap. This change is equivalent to an average increase in immigrants' wages by 3.5%. Both scenarios improve immigrants' labor market outcomes, which might change their marital and spatial distribution. Immigrants, who get richer, become more attractive to natives, and the intermarriage rate might increase. On the other hand, the relative gain from marrying a native decreases from the immigrant perspective, which can lead to the opposite outcome. Similarly, adjustments in labor and marriage market conditions could influence overall region attractiveness and lead to changes in the spatial segregation of immigrants. The direction and magnitude of changes are evaluated using the model and estimated parameters presented in the previous chapter of the paper.

Table 1.10 presents changes in the spatial concentration of immigrants for both scenarios by gender and education level. The spatial concentration is measured as a total variation distance between uniform and observed distribution. Distribution across regions of women and college-educated agents seems to be more concentrated, but the differences are relatively minor. Increasing the wages of immigrants such that the gap between them and natives drop by 25% lead to a decrease in the spatial distribution of all groups. The size of the effect is limited, however, it might improve immigrants' integration. The decline is more substantial among women, especially the college-educated ones. It comes from the fact that the initial wage gap among immigrant women is the biggest, so they respond stronger to changes in its relative value. The decrease in the spatial concentration of immigrants is much stronger in the case of scenario II. The effect is massive among college-educated immigrants and equals around 32% for both genders. It suggests that decreasing the regional wage gap among immigrants has a strong and positive effect on their immigration, measured by spatial concentration.

Table 1.10: Changes in immigrants' spatial concentration

	Baseline	Scenario	
		I	II
		Δ%	Δ%
A. Men			
Noncollege	0.132	-1.7%	-9.3%
College	0.143	-1.7%	-31.7%
Total	0.134	-1.7%	-15.1%
B. Women			
Noncollege	0.146	-2.9%	-11.8%
College	0.147	-8.5%	-31.9%
Total	0.146	-4.0%	-15.9%

Notes: Cells in the table present the spatial concentration of immigrants and its changes under counterfactual scenarios. The spatial concentration of immigrants is measured using the total variation distance between uniform and observed distributions. The sample consists of a cohort of agents born in the '60s.

Analyzed scenarios improve the economic situation of immigrants. Further, it induces changes in marriage market outcomes since matching depends on, among others, economic value. Table 1.11 presents changes to the single rate among immigrant men and women by education level. The first column contains a prediction of the model based on estimated consistent parameters. The remaining columns present the percentage point differences between the baseline for each scenario. Under scenario I (reduction in immigrant-native wage gap), the share of single immigrant decrease in all groups (from -0.2% to -8.2%). The decrease is more substantial among men than among women. It occurs because men work, on average, more, so wage changes impact them to a greater extent. Also, college-educated immigrants react stronger than noncollege ones (-8.2% vs -2% and -1.9% vs -0.2%). The discrepancy, again, is driven by differences in labor supply. Similar patterns can be observed under the scenario I, but the decrease in the single rate is smaller, except for noncollege educate women. A decrease in the single rate suggests that considered labor market changes increase immigrant attractiveness as partners and induce more marriages. Understanding the effect of these changes on social integration requires an analysis of changes in marriage patterns.

Table 1.11: Single rate among immigrants

	Baseline	Scenario	
		I	II
	%	Δ	Δ
A. Men			
Noncollege	18.5%	-2.0%	-1.6%
College	14.8%	-8.2%	-5.3%
Total	17.8%	-3.2%	-2.3%
B. Women			
Noncollege	5.3%	-0.2%	-0.5%
College	9.0%	-1.9%	-1.8%
Total	6.1%	-0.5%	-0.8%

Notes: Cells in the table present the shares of single immigrantsand their percentage point changes under counterfactual scenarios. The share of single immigrants is calculated as the number of singles divided by the total population by gender, origin, and education group. The sample consists of a cohort of agents born in the '60s.

Table 1.12 presents the share of immigrants married to natives by gender and education level. The first column contains model prediction using consistently estimated parameters. The remaining columns present the percentage point differences between the baseline and two counterfactual scenarios. Panel A of Table 1.12 contains those values for immigrant men. Under both scenarios, the share of intermarried male immigrant change in a very similar way. Introduced policies increase the probability of being married to a native. However, the increase in the intermarriage rate is much lower than the increase in the marriage rate showed in Table 1.11, indicating that single immigrants in the baseline scenario still more often marry other immigrants than natives in the alternative scenarios. However, the overall effect of both policies on the social integration of immigrant men is positive.

Panel B of Table 1.12 presents the same values as panel A but for immigrant women instead. In scenario I, the share of intermarried women decreases by 1.9 p.p. for noncollege- and 2.4 pp for college-educated. It occurs because the utility of marriage with male immigrants increases (due to wage growth caused by closing the immigrant-native gap), and they become more attractive partners. The difference between the two education groups comes from the fact that college-educated immigrant women are more likely to be single than noncollege ones in the baseline scenario. It means that the decrease in intermarriage comes not only from women who change their partners' origin. It also comes from women who are single under the baseline scenario but, due to policy changes, not anymore, and they marry immigrant men more often. In the case of scenario II, the magnitude of changes is smaller (-0.5 pp and -1 pp, respectively). However, the relative value and sources of the changes remain the same. It means that both policies negatively impact the social integration of immigrant women.

Table 1.12: Intermarriage rate among immigrants

	Baseline	Scenario	
		I	II
	%	Δ	Δ
A. Men			
Noncollege	28.8%	0.9%	0.9%
College	32.3%	2.2%	2.1%
Total	29.5%	1.1%	1.1%
B. Women			
Noncollege	31.3%	-1.9%	-0.5%
College	32.4%	-2.4%	-1.0%
Total	31.5%	-2.0%	-0.6%

Notes: Cells in the table present the shares of intermarried immigrants and their percentage point changes under counterfactual scenarios. The share of intermarried immigrants is calculated as the number of immigrants married to natives divided by the total population of agents by gender, origin, and education group. The sample consists of a cohort of agents born in the '60s.

Despite the integration aspects, such as the intermarriage rate or spatial segregation of immigrants, policymakers also care about the consequences of their policies on income inequality or welfare. Table 1.13 presents the baseline level of income inequality and its percentage point changes under two scenarios with and without adjustments for location and marriage choices. Income inequality is calculated as a percentage difference between the average per capita income of immigrants and natives by gender and education level. The baseline values are slightly more unfavorable for women than for men. Reducing the immigrant-native wage gap decreases income inequality (from 3.2% to 1.9%) without controlling for adjustments. Even though the average wage increase is higher for women than men due to labor supply choices, the positive effect on income inequality is weaker for women than for men. Further, allowing for adjustments in marriage and region choices leads to a decrease in the effect size. In the case of college-educated men, the primary source of the decrease is that under the counterfactual scenario, they are more often married, so they share the income with partners who are less likely to work. In the case of women, the decrease comes from the fact that under the baseline scenario, they are more often married to natives. The better labor situation of immigrant men makes immigrant women willing to trade higher income for taste in similarity, which mitigates the positive effect on income inequality. The exception to this pattern are noncollege-educated men. In their case, the income inequality gets even smaller while accounting for the marriage market adjustments. It comes from the fact that they are the only group that marries more often with natives, who have higher wages and work more often. Regarding scenario II, the changes in income inequality are smaller but stay positive. They also show similar patterns when controlling or not for adjustments in the marriage market and region choices. Analyses of counterfactual scenarios suggest that both outcomes of labor market integration politics decrease income inequality. However, changes in region and partner choices partially mitigate the positive effect for all groups except noncollege-educated men.

Table 1.13: Changes in income inequality

	Baseline	Scenario I			Scenario II		
		(1)	(2)	(3)	(1)	(2)	(3)
		Δ	Δ	Δ	Δ	Δ	Δ
A. Men							
Noncollege	-12.6%	2.4%	2.7%	2.7%	1.7%	1.9%	1.9%
College	-11.3%	3.2%	2.5%	2.5%	3.0%	2.4%	2.1%
B. Women							
Noncollege	-13.2%	1.9%	1.7%	1.6%	1.4%	1.4%	1.3%
College	-14.5%	3.0%	3.0%	2.9%	2.3%	2.5%	2.0%
Marriage market adj.		✗	✓	✓	✗	✓	✓
Region choice adj.		✗	✗	✓	✗	✗	✓

Notes: Cells in table present income inequality and its percentage point changes under counterfactual scenarios. Income inequality measured as parecentage difference between average per capita income of immigrant and native of the same gender and education level.The sample consists of a cohort of agents born in the '60s.

Table 1.14 presents welfare changes under two policy scenarios. The changes are disaggregated in the same way as in Table 1.13. In scenario I, there is an increase in the welfare of all immigrants (from 1.2% to 3.2%), keeping the baseline distribution of marriages and location choices. Since immigrants' wages rise, households with immigrants have higher disposable income, which leads to utility gains. The increase is more significant for college-educated immigrants than for noncollege-educated ones. The difference is driven by the initial higher wages of better-educated agents. Allowing for adjustment in marital choices increases the welfare of all immigrants except noncollege-educated women. The less positive or negative impact of marriage market adjustment for female welfare emerges from the bargaining between partners. Men work, on average, more than women, so increasing their wages improves their negotiation situation. As a result, they can negotiate more favorable Pareto weights. Endogenizing regional choices do not significantly change welfare outcomes because closing the immigrant-native gap is parallel in all regions. The welfare consequences of reducing regional variation in immigrants' wages are positive for all types of immigrants. The patterns of changes are similar to the one under scenario I, except for positive change induced by adjustment in region choice. The increasing welfare gains associated with adjustment in marriage and location choices contradict rising income inequality shown in Table 1.13. It suggests that even though the income per capita after adjustment decreases, it is compensated by the higher utility of marriage. In particular, agents compensate for the decreasing income by a taste for similarity, marrying more often with other immigrants.

Table 1.14: Welfare changes for immigrants

	Scenario I			Scenario II		
	(1)	(2)	(3)	(1)	(2)	(3)
A. Men						
Noncollege	1.2%	1.8%	1.8%	0.8%	1.1%	1.1%
College	3.2%	3.7%	3.7%	2.4%	2.7%	2.9%
Total	1.7%	2.3%	2.3%	1.3%	1.5%	1.6%
B. Women						
Noncollege	1.5%	1.4%	1.4%	1.1%	1.1%	1.1%
College	2.8%	3.0%	3.0%	2.0%	2.4%	2.5%
Total	1.8%	1.8%	1.8%	1.3%	1.4%	1.5%
Marriage market adj.	✗	✓	✓	✗	✓	✓
Region choice adj.	✗	✗	✓	✗	✗	✓

Notes: Cells in table present percentage changes in welfare of immigrant men and women by human capital.The sample consists of a cohort of agents born in the '60s.

An analysis of the two policy scenarios reveals that changes in labor market conditions lead to an adjustment in marital and spatial distribution. It underlines the importance of including those two aspects in the analysis of outcomes of labor market integration policies. First, analyzed policies decrease the spatial concentration of immigrants, which could strengthen the positive effect of labor market changes for integration. Second, both policies increase the number of marriages and intermarriages among immigrant men. They also increase the number of marriages among immigrant women but decrease intermarriages among them. It suggests that improving immigrants' labor market conditions differently impacts the social integration of men and women via marriages with natives. Third, income inequality increases while I control for adjustment in marriage and location choices. It suggests that immigrants trade economic gains for better region or marriage perspectives. Finally, counterfactual scenarios showed that improving the labor market situation of immigrants allows for obtaining additional welfare gains through marital and regional sorting changes.

1.6. Conclusion

In this paper, I estimate a three-stage structural model using German micro-data to quantify the effect of labor market integration policies on intermarriage and spatial concentration of immigrants. Further, I evaluate how those changes impact immigrant-native income inequality and immigrants' welfare. Estimated parameters suggest that immigrants, on average, earn less than natives of the same gender and education. It makes them less attractive partners and, together with estimated strong preferences for similarity, leads to a relatively low number of intermarriages. I use the model to simulate the effects of reducing the immigrant-native wage gap by 25% and decreasing regional variation in immigrants' wages by 50%. These exercises lead to three main conclusions. First, closing the immigrant-native wage gap positively affects the spatial concentration of immigrants because regions with a more significant gap (and, at the same time, smaller migration populations) become more attractive. Further, the policy's impact on the frequency of marriages with natives varies by immigrant gender. The increase in wages stronger affects the economic attractiveness of immigrant men since they have a higher labor supply. As a result, the policy leads to more marriages between them and native women. On the other hand, due to the gender difference in the size of the policy's effect, immigrant women marry immigrants more often. As a result, the changes in homogamy differently affect the social integration of immigrant men and women. Second, reducing regional variation in immigrants' wages significantly changes their regional distribution. Immigrants move out from the South (the region offering the highest wages in the baseline scenario) to settle in the North and West. It significantly flattens the distribution and leads to a decrease in spatial concentration across genders and education levels up to 32%. This finding suggests that reducing regional wage variation by increasing immigrants' earnings in regions initially characterized by lower wages might be a powerful policy tool. Not only it decreases income inequalities, but also positively impacts integration via a decline in spatial concentration. However, these positive boosts are partially mitigated by the decrease in the intermarriage rate among immigrant women. Finally, both policy scenarios improve welfare levels and lower the income inequality between immigrants and natives. Adjustments in marriage and location choices lead to, on average, a simultaneous decrease in the drop of income distribution and an increase in welfare gains. It suggests that after improvement in labor market conditions, immigrants trade part of the gains for noneconomic gains associated with marriage and location choices.

2. PARENTING STYLES AND CHILDREN'S SKILL DEVELOPMENT

2.1. Introduction

Parenting decisions, including parenting style, shape children's skills early on in life and influence their long-run accumulation of human capital. Human capital, represented by cognitive and non-cognitive skills, is a crucial determinant of individuals' well-being, wages, and health (Conti et al. (2019), Hanushek and Woessmann (2008)). Children from lower socioeconomic backgrounds start to display lower skill levels than their peers early in life, which is widely documented across contexts (Cunha et al. (2006), Heckman and Mosso (2014), Attanasio et al. (2020a)). The gap emerges in cognitive skills like mathematical reasoning, logical thinking, or language skills, and non-cognitive or socio-emotional skills, which influence how individuals interact with others and navigate social situations. In the long run, these differential development patterns result in lower intergenerational mobility and higher inequality (Attanasio et al. (2020c)).

Numerous policies aim to close the skill gap, but to do so, the main drivers behind the skill gap need to be identified. To design effective policies to close the gap, it is crucial to understand when it emerges and in which behaviors and conditions it originates. While factors such as lower investments, initial skills, and, e.g., peers are well documented (see Heckman and Mosso (2014) and Almond et al. (2018) for an overview), less is known about the impact of parenting style. Parents' overall approach and pattern of behaviors when interacting with their children can influence their skill development (Doepke and Zilibotti (2017)). How parents establish rules, offer guidance, and respond to their children's needs, behaviors, and emotions might influence children's cognitive and non-cognitive skill development. Child development policies like parenting training interventions have been increasingly used to improve parenting skills. They aim to enhance skill development and close the skill gap between children from different socioeconomic backgrounds. However, to design these interventions more effectively, decision-makers need to know which parenting styles positively impact skill development, which parenting behaviors to target, and at which stages of childhood they are most important.

Answering these questions is challenging, as parents' parenting style might correlate with other characteristics, e.g., initial ability, which influence children's skill development. Additionally, identifying the impact of parenting skills on a child's skill development requires a long-run panel data structure. Surveys collecting detailed data on parenting style dimensions, capturing, e.g., how consistent parents are and how much they explain rules to the children, are rare. In general, the literature has focused so far mainly on other parental investments, such as time or monetary investments (Cunha et al. (2010), Del Boca et al. (2014), Caucutt et al. (2020)). Some papers include parenting style as a factor in skill development but focus on a particular dimension, e.g., if parents interfere with the choice of friends of their children (Agostinelli et al. (2023)). Others like Del Bono et al. (2016), Del Boca et al. (2016), Fiorini and Keane (2014), Le Forner (2021) do include broader definitions of parenting style in their estimations of skill production functions. However, they focus on other factors like time investments in the analysis. Therefore, they use a reduced number of parenting dimensions and, due to that, do not explore its multi-dimensionality. However, the effective design of policy interventions requires a better understanding of which parenting dimensions are most influential. If parental warmth is a key determinant, interventions should aim to increase warm behaviours of parents, as hugs for example. Additionally, to model parenting style in skill development, one needs to know if it enters the production function as an input itself or if it rather enhances the productivity of other investments. Finally, extreme parenting styles might disproportionately impact the children's skills, which would suggest a non-linear relationship between those two factors.

Therefore, in this paper we investigate how parenting style influences human capital development in middle childhood and adolescence. We estimate the impact of different parenting dimensions on cognitive and non-cognitive skill development of children. Doing so, we study how parenting style influences skill development and which dimensions of parenting style influence skill development most. Further, we investigate if these relationships are non-linear and test if parenting style influences the productivity of other inputs as time investments (time spend with the child by parents). The analysis provided can help to design parenting training interventions more effectively and be informative for design of models for children's skill formation.

In order to accomplish this, we use data from Australia, the Longitudinal Study of Australian Children (LSAC), supplied by the Australian Government Department of Social Services in collaboration with the Australian Institute of Family Studies. This panel dataset contains observations of approximately 10,000 children from two birth cohorts: one followed from age 0-1 to 14-15 (younger cohort) and the other from age 4-5 to 18-19 (older cohort). The LSAC dataset offers comprehensive longitudinal information on parenting styles and other relevant factors such as time investments and measures of children's skills. Given data availability, we focus on the age range of 8-15.

To measure parenting style, we employ a factor analysis on various survey modules targeting parenting. This approach yields five distinct dimensions of parenting style. Firstly, parental warmth captures how much affection parents express to their children. Secondly, parental reasoning assesses how parents explain rules and consequences to their children. The third dimension, parental hostility, captures how often parents praise the child for positive behavior or react angrily in response to negative behavior. The fourth dimension attempted consistency, evaluates how often parents attempt to reinforce the completion of requests and punishments for non-compliance. Lastly, the fifth dimension, inconsistency, measures how often the child gets out of such punishment. We use Matrix Reasoning (MRT) and Peabody Picture Vocabulary (PPTV) tests to measure cognitive skills. Strength and Difficulties Questionnaires (SDQ) measure non-cognitive skills, which record behavioral, emotional, and conduct-related problems.

To estimate the influence of parenting dimensions on skill development, we exploit the panel structure of the LSAC dataset and the comprehensive range of controls. This approach allows us to tackle potential issues such as unobserved ability and endogeneity. However, it is important to acknowledge that none of the employed models can entirely eliminate all potential biases, given the absence of exogenous variation. Therefore, we carefully analyze the outcomes of each econometric specification to assess if they consistently indicate a significant impact of parenting style on skill development.

In addition, we run several robustness checks to account for different ways of creating parenting style measures and include varying sets of controls. These checks serve to assess the stability and reliability of our estimates. Thanks to that, our analysis can enrich the existing literature and offer valuable insights for future policy interventions or randomized control trials to test these findings further. They might also inform modeling choices for skill formation.

We find a significant negative impact of parental hostility on the development of non-cognitive skills. Specifically, a one standard deviation increase in hostility corresponds to a decrease in non-cognitive skills ranging from 0.12 to 0.35 standard deviations. This finding is consistent across age groups, with increasing magnitudes as children grow older. We observe a comparatively weaker but still negative influence of inconsistency and parental reasoning on non-cognitive skill development. On the other hand, parental warmth leads to small increases in non-cognitive skills. Interestingly, our results indicate that high levels of hostility exert an even greater negative impact on skills, suggesting a non-linear relationship. Notably, we do not find parenting style to influence the productivity of time investments in skills formation. Thus, we can conclude that incorporating parenting style as an additional input in the skill production function is a more accurate modeling approach.

Our study indicates that the connection between authoritative parenting (low hostility, high warmth, consistency, and reasoning) and non-cognitive skill development, as documented in previous research (see Spera (2005) and Doepke and Zilibotti (2019) for an overview), is primarily influenced by low levels of hostility and inconsistency. Parenting warmth and reasoning have only a minor impact. In contrast, for authoritarian parents, the beneficial effect of consistency is outweighed by the negative impact of higher levels of hostility. We additionally document, that parents with lower income display higher levels of hostile and inconsistent parenting, which may contribute to the skills gap between children from different socioeconomic backgrounds in the case of non-cognitive skills. For example, hostile and inconsistent parenting can arise from stress (Sanders and Woolley (2005), Bloomfield and Kendall (2012) and Hutchison et al. (2016)), which parents with lower income experience to a higher level. As we do not find a parental style to influence cognitive skills, other factors like time investments or school environment might play a bigger role in the development process of these skills.

Our findings show that parental training programs aimed at reducing hostility might be more effective than programs targeting other dimensions of parenting, assuming that parental behavior is equally amenable across dimensions. The results indicate a progressively stronger negative effect on non-cognitive skills with increasing levels of hostility. Therefore, focusing on households where hostile parenting is prevalent may be the most efficient approach to increasing non-cognitive skills. In contrast, we do not find consistent impacts of any parenting dimension on cognitive skills. These findings suggest that if the goal is to increase cognitive skills solely, policy interventions should consider targeting other factors beyond parenting. However, a combined approach may be necessary to increase the overall skill level.

Related literature: As we look at human capital development and how to improve it, our paper links to the literature on skill development. Models as Cunha and Heckman (2008), Cunha et al. (2010), Attanasio et al. (2020c) Attanasio et al. (2020b) study how children's skills dynamically accumulate over time along the dimension of health, cognition, and socio-emotional (non-cognitive) skills. Enriching this process with endogenous parental investment choices, Todd and Wolpin (2007), Del Boca et al. (2014), Lee and Seshadri (2019), Caucutt et al. (2020), Wiswall and Agostinelli (2020) model how parents decisions influence their children's outcomes. In this context, the listed models abstract from modeling the influence of parenting style and proxy it by parental background characteristics or modeling it as unobserved heterogeneity.

Human capital development models, including parenting style as an additional input in skill formation, are, for instance, Lizzeri and Siniscalchi (2008), Cunha (2015), Del Boca et al. (2016), Cobb-Clark et al. (2019), Kim (2019) and Falk et al. (2021). These papers build a theoretical framework for including parenting style in the process and empirically support underlying assumptions. They propose to include parenting style as an additional input in the production function, as Dooley and Stewart (2007), Fiorini and Keane (2014), Del Boca et al. (2016), Kim (2019) and Falk et al. (2021) empirically show that different styles can have impacts on skill development (mainly non-cognitive skills, but also to some extend educational achievement). Related to Dooley and Stewart (2007) and Kim (2019), we, in particular, analyze the negative effects of hostile parenting/punishment (e.g., angrily shouting), but take a step further and look at these factors in interaction with other parental investments and at different ages to get a better understanding of the skill production function.

Regarding the increasing use of parenting training in combination with early childhood interventions, a deeper understanding of how parental style dimensions influence investment decisions and skill development could give insights into how intervention can be designed more effectively. If a particular dimension of parenting style has a large impact on skill development, it might be most effective to target that dimension if it is malleable. Our analysis aims to enrich the literature by giving new insights on how to model parenting style in children's skill formation process and potentially design interventions more effectively.

Doing so, we link to the literature on parenting style in economics and developmental psychology. In economics, the literature has focused on the impact of different parenting styles following Baumrind (1967) and Maccoby and Martin (1983) like permissive, neglecting, authoritarian, and authoritative style on skill development (see Doepke et al. (2019) for an overview). The styles summarize the extent to which parents choose to intervene in their children's behavior. For instance, see Doepke and Zilibotti (2017) who define the following: parents exert a permissive style when they leave children their independence and are supportive but not strict. This is contrasted by an authoritarian style, where parents impose their will through coercion strictly and are not supportive. Parents can instead also be authoritative; which is when they aim to affect the child's choice using persuasion and are strict but supportive. Another category are neglectful parents who are neither strict nor supportive.

Doepke and Zilibotti (2017) find an association between higher educational outcomes and authoritative and, somewhat less extent, permissive parenting compared to neglecting and authoritarian parenting using US data. Distinguishing between authoritarian and non-authoritarian styles of intervening with peer interactions of adolescence, Agostinelli et al. (2023) find that positive impacts of interventions like moving children to better neighborhood are smaller as parents push back on children's new peer groups (if they have an authoritarian styles). This result highlights the importance of investigating the interaction between parenting style and investments of parents as well as children's environment. In contrast to these papers, we do not focus on parenting style as a choice to which extent parents influence children's behaviors. Instead, we would like to determine the components of parenting styles that influence skill development most and should be the target of parenting training. Thus, we study separately parenting dimensions such as parental warmth, reasoning, consistency, and hostility.

This strategy links us to the literature on developmental psychology, which studies parenting styles and their impact on skill development. Also, this literature defines styles using Baumrind (1967)'s categories. However, following Spera (2005) and McWhirter et al. (2023) styles vary in definition slightly from the economics literature. Authoritarian style is described as low in warmth and responsiveness; parents are strict and demanding, expecting obedience, and do not reason for rules. They assert power and use punishment if a child misbehaves and score high on control. The authoritative style is characterized by warmth, responsiveness, high reasoning, demandingness, and scoring high on control. In contrast, neglecting/indulgent parents score low on responsiveness, warmth, and control, and permissive ones moderately in responsiveness, low on control, and high on warmth while they are not demanding.

The literature finds authoritarian and neglecting styles negatively associated with non-cognitive skills. In contrast, the authoritative style has a positive association, confirming

Doepke and Zilibotti (2017)'s results (see Spera (2005), Fletcher et al. (2008), García and Gracia (2009), Luyckx et al. (2011), Howenstein et al. (2015) and McWhirter et al. (2023)). However, most papers suffer from small sample sizes and do not use panel data. By exploiting a large longitudinal survey, we can leverage the panel data structure to provide a more structured analysis of different components of parenting style. Doing so allows us to correct for unobserved factors which could confound the analysis, like a parents' selection of a certain parenting style due to their initial ability, which also impacts children's skill outcomes. To enrich the existing literature, we also look at the interaction of parental investments with parenting styles and the potentially non-linear relationship between parenting styles and children's skills.

Looking at the context of Australia, we analyze the impact of different dimensions of parenting style and their interaction with investments in children's skill development in high-income countries. Other papers have studied skill development in this context using the LSAC data[1]. Fiorini and Keane (2014) and Le Forner (2021) focus on the impact of time investments on children's skills, similar to the analysis of Del Bono et al. (2016) for the UK. Summarizing different parenting style components using principal component analysis, they find parental warmth and authoritarian style to influence non-cognitive skills, while time investments do not. Building on their results, we complement the literature by analyzing different components of parenting style, the interaction with investments, and a larger sample. Additionally, we extend the analysis by looking at a longer period, spanning middle childhood, an often understudied period in child development (see Almond et al. (2018)). This could give additional insights on when to best implement parenting training interventions and when their impact is best measured.

The rest of the paper is organized as follows. In Section 2.2, we describe the data used and present relevant empirical facts on parenting skills and skill development in Australia. Next, we introduce the empirical framework in Section 2.3. In Section 2.4, we discuss results, followed by concluding remarks and ideas for future research in Section 2.5.

2.2. Data

2.2.1. Data sources and construction

The Longitudinal Study of Australian Children (LSAC) is a biannual survey following two cohorts of Australian children since 2004. The older cohort ("K cohort") was born between March 1999 and February 2000 (4,983 children), and is followed from age 4-5 to

[1]Australian data in child development, as the LSAC and LSIC (Longitudinal study of Indigenous children) data sets have been used to study in particular non-cognitive skill development due to the richness of their measures. For instance, Guy et al. (2016) and Twizeyemariya et al. (2017) study the occurrence of mental health risks for Australian children, while Christensen et al. (2017) study the impact of these risk factors on non-cognitive skills over time.

18-19. The younger cohort (”B cohort”) is born between March 2003 and February 2004 (5,107 children) and followed from age 0-1 to 14-15. Both cohorts were surveyed biannually from 2004 to 2020. The survey collects information about the children and their parents, along with measures of child development, including cognitive and non-cognitive skills. The advantage of the LSAC data set is that it combines detailed information on parenting styles with time-use diaries and children's skills and demographics. This feature allows a rigorous analysis of the impact of parental styles on children's skill outcomes taking into account other parental investments like time spent with the child. Additionally, interactions between parenting styles and time investments can be investigated, and the impact on different types of non-cognitive skills can be compared across different ages. In particular, the richness of the parenting style questions allows us to explore different dimensions of parenting style and their impact.

For the analysis, we pool both cohorts together and compare their outcomes at the same ages. Therefore, we estimate impacts in age groups consisting of two years: ages 8-9, 10-11, 12-13, 14-15, and 16-17. As the survey is conducted biannually, this is the most granular level possible. In our main analysis, we exclude outliers and restrict the sample to observations with available skill measures, parenting style, time investments, and necessary control variables.

Non-cognitive skills

The LSAC measures non-cognitive skills by a strength and difficulties questionnaire (SDQ) filled out by parents for children aged 6-15.[2] The SDQ consists of 25 questions covering five subjects: emotional health, behavioral problems (conduct), hyperactivity issues, peer problems, and pro-social behavior. Following Goodman et al. (2010) and Le Forner (2021), one can summarize these subjects into four broader indexes. The indexes are emotional skills (internalizing SDQ), behavioral skills (externalizing SDQ), and pro-social skills (social SDQ). Behavioral skills capture behavioral problems and hyperactivity issues. In contrast, emotional skills entail questions about emotional health and peer problems. Finally, pro-social skills are the index for pro-social behavior. Behavioral and emotional skills can be summarized to the total SDQ as an index.

To keep results tractable, we restrict our analysis for now to one index and follow Le Forner (2021) in using the total SDQ, which is the sum of behavioral and emotional skills. By doing so, we intend to capture non-cognitive skills in their various dimensions in one index. We standardize this measure by age group to facilitate the interpretation of estimated coefficients and comparability across different age groups.

[2] The survey also collect information from teachers. We use only the parental assessments in the analysis. The teacher questionnaires suffer from missing information, which is why we abstract from using them.

Cognitive skills

We use two measures for cognitive skills which are available in the LSAC. Firstly, the Peabody Picture Vocabulary Test (PPVT) which measures children's knowledge of the meaning of spoken words and their receptive vocabulary. The PPVT is adjusted to age in terms of difficulty and administered in the survey for children aged 4-5 years, 6-7 years, and 8-9 years. To conduct the test, children are shown 40 plates of pictures in a PPVT stimuli book and told a word to which they were required to choose the picture which best represents the meaning of the word. They could do so by pointing a picture or saying the number of a picture. Test scores are calculated using Rasch Modelling to ensure changes in scores relate to real changes in knowledge not changes in position relative to peers. As the test is only administered up to age 9, we cannot compare the outcomes to older age groups.

The second measure available in the LSAC is a Matrix Reasoning Test (MRT). This test was administered to children at ages 6-7 years, 8-9 years and 10-11 years. The test is a nonverbal intelligence test consisting of 35 items of increasing complexity. Each item is an incomplete set of diagrams, and the child is required to complete the set from five different options. The test score is the number of correct responses, the child gave, scaled based on age norms (determined in the WISC-IV manual). As the test is only administered up to age 11, we cannot compare the outcomes to older age groups. Therefore, we mainly focus on non-cognitive skills when comparing coefficients over childhood. As there are no measures in the survey beyond age 11, we have to restrict the analysis to that age group for cognitive skills. We standardize both cognitive measures by age for comparability and interpretation facility.

Parenting style

The survey collects information on both parents and their behaviour towards the child. Parenting questions are consistently asked across waves in four different areas: hostile parenting, parental warmth, consistent parenting and inductive reasoning (see appendix Table B.1 for a more detailed description of each subgroup). We abstract from using other available information, which is not consistently available across waves to avoid not comparable measures and use the information available for the principal care giver. Following Fiorini and Keane (2014), Del Bono et al. (2016) and Le Forner (2021), we use factor analysis to derive dimensions of parenting style using the survey information. Given, that we conduct a detailed analysis of parenting style, we keep as many dimensions as possible. In contrast to Fiorini and Keane (2014) and Le Forner (2021), who also use the LSAC data, we do not pool parenting questions together to get as least factors as possible. As our paper focuses on the impact of parenting style on skill development and not on time investments, we investigate how the different components influence skill development to isolate which dimension to target in parenting training. Therefore, we conduct a separate factor analysis for each of the four areas as we would like to analyse each dimension of parenting style. We conduct this analysis wave by wave. We retain factors with eigenvalues larger than 1 and factors are rotated. As for inductive reasoning, at age 4-5 and age 6-7 only 2 respectively 3 of the 5 questions asked in other waves are included, we use only those available in those waves.

Table B.2 shows the rotated factor loading coefficients of the principal component analysis for each measure and each wave. Factor loadings which are larger than 0.25 in absolute value are displayed in bold. The principal component analysis for each measure leads to one factor pooling all sub questions (eigenvalues¿1). Only for the measure consistency two factors are needed to summarize the variation. The first factor can be described as inconsistent parenting style, the child gets out of punishment or ignores it. The second factor captures if parents attempt to make the child fulfill requests and attempt to punish it if not. We will call this factor: attempted consistency. Factor loadings are stable across waves, except for parental consistency in wave 3, here only one factor is needed to describe the variation (inconsistency). Hence, we conclude that measures are comparable across waves despite for consistency in wave 3.

Later in the analysis, we interact parenting style with investments to determine if parenting style influences the impact of investments and if there is a quality-quantity trade-off. To do so, we summarize the variation in dimensions of parenting style in additional factor analysis in the style of Diamond (2016). We do so to minimize the number of necessary interactions. Doing so, we proceed as before; results of the factor analysis are displayed in Table B.3. We find that two factors are needed to summarize the variation in parenting styles (except for wave 3 with one). Factor one can be described as loading on parental warmth, reasoning, and attempted consistency. Factor two loads on hostility and inconsistency. From wave 6, the assignment of these factors switches, so factor two loads on parental warmth, reasoning, and attempted consistency, and factor one on hostility and inconsistency. To keep consistency across waves, we frame factor 1 for the age group younger than 14 and factor 2 for the age group older than 14 as an empathetic parenting style (warm, reasoning, and attempted consistency). Factor 2 for the age group younger than 14 and factor 1 for the age group older than 14 are framed as harsh parenting (hostile and inconsistent). Similarly to the dimensions of parenting styles, the loadings for parenting styles are fairly consistent across waves except for wave 3.

Time investments

In the LSAC dataset, Time Use Diaries (TUD) are utilized to gather data on children's activities. The data collection process involves two methods. For cohort K, spanning three waves (ages 4-9), and cohort B, also across three waves (ages 0-5), data is collected over two 24-hour periods, typically one on a weekday and another on a weekend day. The information is recorded on paper diaries, divided into 96 15-minute intervals, which parents fill out. Parents select the activity, location, and individuals involved from a predefined set of options.

For cohort K, spanning three waves (ages 10-15), and cohort B, also across three waves (ages 10-15), children themselves become the informants (with support from the interviewer). Furthermore, the Time Use Diaries undergo significant changes. Instead of paper diaries, data is now collected using a computer instrument. Additionally, the time span of activities is not limited to 15-minute intervals. Moreover, activities are recorded only on a single day of the week, either a weekday or a weekend day. However, similar to the previous version, children complete the diary by selecting the activity, location, and individuals involved from a predetermined set of options.

To analyze the effect of parental time investment on children's cognitive and non-cognitive skills, we aggregate the recorded activities into five main groups:

1. Educational activities with parents
2. Educational activities with adults other than parents
3. General care with parents
4. General care with adults other than parents
5. Other time

Since time investments are not the focus of our analysis, we follow the aggregation rules established by Fiorini and Keane (2014) and Le Forner (2021) to group activities. It is important to note that the set of alternatives may change over time, but the primary divisions between educational, general care, and other activities remain consistent across different survey waves. In cases where multiple activities are reported simultaneously, we prioritize the primary activity. If information about the activity is missing, we assign it to the category other time. This ensures that the total time spent on activities always sums up to 24 hours. Regarding time spent with adults other than parents, we only consider it if the activity was conducted with adults while parents were not participating. If parents were involved in the activity, it is classified as time spent with parents.

The two methods of TUD collection differ in terms of the days of collection. For these waves, when TUDs were collected on both weekdays and weekends, we calculate a weighted average for each time input. Weekdays are assigned a weight of 5, while weekend days are assigned a weight of 2. However, for the remaining waves, data collection was conducted on a single day only. As a result, in all regressions, we include dummy variables to indicate whether the record was on a weekday, weekend day, or an average of both diaries.

2.2.2. Demographics

Table 2.1 shows statistics for main descriptive characteristics of the sample population across different age groups. The age groups range from 4-5 years to 14-15 years. In terms of child characteristics, the table shows that approximately half of the children are female. The percentage of indigenous children is relatively low and stays constant across age groups. The percentage of children living with both parents decreases with age, because parents may separate or divorce as the child grows older, leading to the child living with only one parent or transitioning between households. Additionally, around half of the children belong to cohort K.

The proportion of primary caregivers with a college education increases slightly over time, which suggest that parents with higher education tend to drop less from the sample. However, it might be also that some parents acquire higher education in the process. The household characteristics indicate that the average number of children under 18 in the household ranges from 1.51 to 1.66. The vast majority of households lives in urban areas.

Table 2.1: Demographic characteristics of the sample

	Age					
	4-5	6-7	8-9	10-11	12-13	14-15
Child:						
Gender	0.49	0.49	0.49	0.49	0.48	0.49
Age	4.22	6.32	8.34	10.38	12.45	14.38
Indigenous	0.04	0.04	0.04	0.04	0.03	0.03
Living with both parents	0.82	0.79	0.76	0.73	0.72	0.69
Born early	0.07	0.07	0.07	0.07	0.07	0.07
Older cohort (K)	0.51	0.51	0.51	0.51	0.51	0.51
Primary caregiver:						
Age	34.83	36.95	39.04	41.10	43.27	45.31
College education	0.27	0.27	0.27	0.27	0.29	0.29
Household:						
Number of children	1.51	1.61	1.66	1.65	1.63	1.55
Weekly income (in AUD)	1,486	1,667	1,918	2,027	2,214	2,257
Urban	0.87	0.87	0.86	0.86	0.86	0.87
Observations	9,285	8,632	8,343	7,858	7,215	6,607

Notes: All means calculated using population weights.

2.2.3. Facts on parenting styles and income

Composition of parenting styles and their association with skills

Given the survey information and results of the principal component analysis, we can look at several dimensions of parenting: parental warmth, hostility, attempted consistency, actual inconsistency and reasoning. Parenting warmth expresses how much parents hug their child, show affection and feel close to the child. Parental hostility describes if the parents rarely praise the child and often disapprove their behaviour, react angry while punishing them. Parents scoring high on attempted consistency often make sure the child completes their requests and punish the child if they do not. Inconsistency then describes, how effective parents are in punishing, thus, if the child gets away from punishment. High values mean parents are not often enforcing their punishments. Lastly, reasoning captures how often parents explain rules and the consequences of the child's behaviour (see Table B.1 for details).

As the literature tends to summarize these dimensions into parenting styles (patterns occurring across parents), we look at their correlation in Table 2.2. Parental warmth and parental hostility are negatively correlated, while warmth positively correlates with reasoning. Attempted consistency is also positively correlated with reasoning, but the magnitude of the correlation coefficient is smaller. In contrast, hostility is positively correlated with inconsistency. Other correlation coefficients are relatively small. By construction, attempted consistency and inconsistency are not correlated, as they originate from the same factor analysis. Overall, the correlations are not very high, suggesting the multi-dimensional character of parenting styles.

Table 2.2: Correlation between parenting dimensions

	Parental warmth	Parental reasoning	Hostile parenting	Attempted consistency	Inconsistent parenting
Parental warmth	1.000				
Parental reasoning	**0.492**	1.000			
Hostile parenting	**-0.385**	-0.047	1.000		
Attempted consistency	0.128	**0.310**	0.036	1.000	
Inconsistent parenting	-0.134	-0.031	**0.442**	0.000	1.000

Notes: Displayed are correlation between different dimensions of parenting styles in the data (exemplary for age group 8-9). Statistics are calculated using population weights.

To compare these dimensions and their correlations with the parenting styles in the literature, we classify the dimensions into styles following Baumrind (1967) and Maccoby and Martin (1983) (see Spera (2005) and McWhirter et al. (2023) for an overview). This classification encompasses four styles:

1. Authoritarian: low warmth and reasoning, high consistency and hostility
2. Authoritative: high warmth, reasoning, and consistency, low hostility
3. Permissive: high warmth, low consistency, and hostility
4. Neglecting: low warmth, reasoning, consistency, and hostility

Regarding the described correlations, the positive association of parental warmth and reasoning indicates patterns of an authoritative parenting style. High hostility would indicate an authoritarian parenting style; however, low consistency does not apply to that and could rather speak of a neglecting style. Let us compare these to the factor analysis summarizing dimensions into parenting styles. We mainly find the variation to describe an authoritative parenting style for the first factor (see Table B.3). Values are high for parents loading on warmth, reasoning, and attempted consistency. Parents scoring low on this factor could be described as neglecting. In contrast, the second-factor loads on hostility and actual inconsistency, which could speak for an authoritarian style. Permissive would be described by loading on the first and second factors jointly, offsetting the hostility in the second factor.

We investigate parenting style dimensions since they are associated with skill outcomes. In our data, this association mainly holds for non-cognitive skills rather than for cognitive ones. We present scatter plots with fitted lines for non-cognitive skills and each dimension of parenting style in Figure 2.1 and for cognitive skills in Figure B.1. Visibly, non-cognitive skills positively correlate with parental warmth and negatively with hostility and inconsistency. For cognitive skills, there is a weak negative correlation with inconsistency. These patterns hint at which factors might parenting training should particularly target.

However, it is important to consider that other factors may also drive these associations. Therefore, our empirical strategy aims to establish a more structured and informative relationship. For instance, the income and education levels of the parents could influence their parenting style. Financial stress, for example, may lead to increased hostility or inconsistency as parents may lack the time, patience, or capacity to enforce rules in a non-angry manner. Additionally, the number of siblings or the gender of the child could act as confounding factors. By accounting for these potential confounders and employing a rigorous empirical approach, we can better understand the nuanced relationship between parenting style dimensions and skill outcomes.

Parenting styles and income

Children from lower socioeconomic backgrounds display lower skill levels than their peers (Cunha et al. (2006), Heckman and Mosso (2014), Attanasio et al. (2020a)). This skill gap is widely documented across contexts and applies to cognitive and non-cognitive skills. Numerous policies aim to close it, but to do so the main drivers behind this gaps need to be identified. To design effective policies to close the gap, it is crucial to understand when it emerges and in which behaviours and conditions it originates in. While factors as lower investments, initial skills and e.g. peers are well documented (see Heckman and Mosso (2014) and Almond et al. (2018) for an overview), less is known on the impact of parenting style. For instance, Cobb-Clark et al. (2019) find a monitoring parenting style, so for instance knowing where the child goes after school, to be negatively correlated

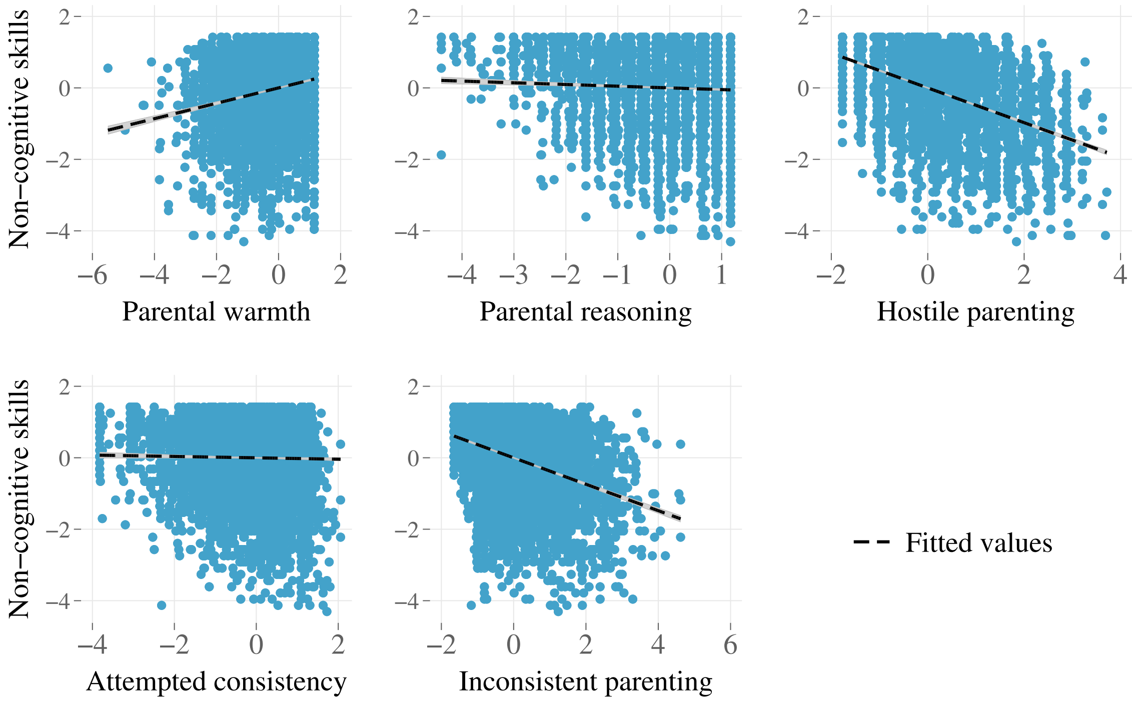


Figure 2.1: Correlation of parenting dimensions with non-cognitive skills

Notes: The figure displays the relationship between non-cognitive skills (measured by the SDQ test) and different parenting styles. Each data point represents a child from the 8-9 age group. In addition to the data points, a line is plotted on the graph, which represents the fitted values based on a linear regression analysis. The line slope is estimated using population weights.

with socioeconomic disadvantage. If parenting styles vary systematically by income or education, it might contribute to the skill gap and be a driver for inequality in children's skills. For example, hostile and inconsistent parenting can arise from stress (Sanders and Woolley (2005), Bloomfield and Kendall (2012) and Hutchison et al. (2016)), which parents with lower income experience to a higher level. Therefore, they might have an on average higher score on hostility, which is negatively associated with non-cognitive skills.

We examine if the distribution of parenting styles in Australia varies by income and education. To do so, we estimate the kernel density of parenting dimensions for different household income groups (the 1st, 3rd, and 5th quintiles) and the primary care gives education level (college and non-college). Figures 2.2 and B.2 illustrate these distributions. Regarding Figure 2.2, our analysis reveals notable differences in parenting styles across income quintiles. Parents in the lower segment of the income distribution are more likely to display high hostility and inconsistency compared to their counterparts in higher income quintiles. Simultaneously, parents with lower income on average show lower warmth, reasoning, and attempted consistency levels. Furthermore, the distributions of parenting dimensions for parents in the 3rd and 5th income quintiles are pretty similar. However, parents in the 3rd quintile demonstrate a slightly lower tendency to persist in establishing consistency. They also have a higher likelihood of exhibiting inconsistency.

Based on Figure 2.1, hostility and inconsistency are associated negatively with non-cognitive skills. Moreover, parenting styles are correlated with income. Thus, the difference in parenting styles might drive a part of the skill gap between children from different socioeconomic backgrounds. As the distribution of parental consistency measures varies the most by income, these might be a parenting dimension that mainly contributes to the observed skill gap between the bottom and the top parts of the income distribution. To investigate this, one needs to control for selection into parenting style and other confounding factors such as parental investment decisions.

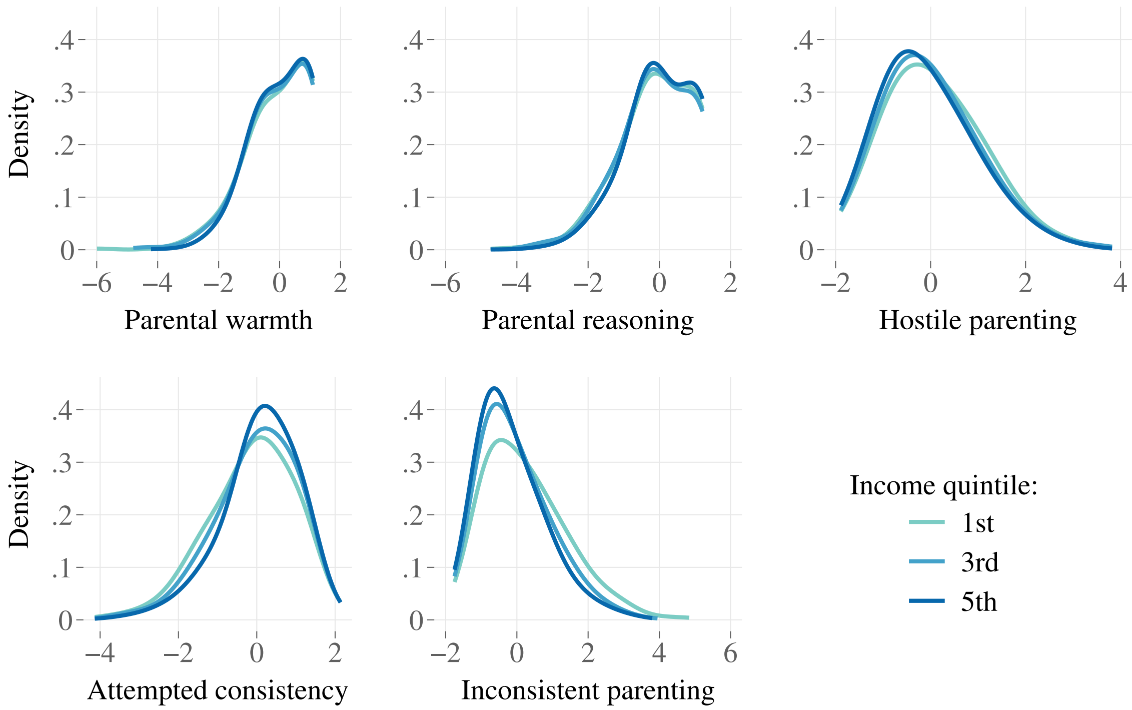


Figure 2.2: Distribution of parenting dimensions by household income

Notes: The figure displays the empirical distribution (smoothed using the kernel function approach with population weights) of different parenting styles by income quintile for children aged 8-9.

Turning to another socioeconomic status factor, primary caregivers' education, the variation in parenting skills is lower. Parents mainly vary in their level of consistency by education (see Figure B.2). Primary caregivers with college degrees are likelier to have consistent parenting levels than those without college. College-educated parents tend to display higher warmth and reasoning. Differences in hostility are minimal. These factors could contribute to children with lower socioeconomic status lagging in skills to their peers, especially regarding non-cognitive skills. Nonetheless, these are pure correlations, and we will use our estimation strategy to disentangle the effects.

2.3. Empirical framework

In this section, we outline the empirical strategy employed to estimate the impact of parenting styles on cognitive and non-cognitive skills. First, we assume the production function of children's skills to take the following form:

$$Y_{ia} = F_a(Z_{ia}, Y_{i0}) + \epsilon_{ia} \tag{2.1}$$

where Y_{ia} is a skill measure for child i at age a, and F_a if an age-specific function transforming production inputs Z_{ia} and the measure of child's initial skills edowement Y_{ia} into the skill level at age a. Production inputs Z_{ia} entail a vector describing past and current parenting style dimensions PS, time investments TI (educational and care time spent with parents and others, and another time spent (i.e., on sleep or socializing), and other parental and household characteristics X up to age a which influence skill development.

To estimate the production function expressed in Equation 2.1 we use the approach of Todd and Wolpin (2003, 2007), applied by Aurino et al. (2019), Fiorini and Keane (2014), Del Bono et al. (2016) and Le Forner (2021). In contrast to these papers, we focus on the impact of parenting style dimensions, rather than time investments or food insecurity, to determine the impact on skill development. Estimating Equation 2.1 without controlling for all inputs and initial endowment can lead to biased estimates due to endogeneity and selection. There are three sources of bias. Firstly, omitted variables can lead to biased estimates if correlated with independent variables. An example could be omitting past investments, which could be related to current ones, and both matter for skill development. Additionally, parents might select into certain parenting style (e.g., driven by education), which directly influences the child's skills and the choice of parenting style. Then, without controlling for education, the coefficient of parenting styles might be biased. Secondly, reversed causality might play a role. Parents might adjust their investments due to skill outcomes, compensating, for example, low skills with higher investments. Thirdly, measurement errors in skills and investments can bias results.

Each of the specifications we employ deals with some of these biases. While presenting different specifications, we thoroughly discuss their results and limitations to draw conclusions on underlying relationships. However, given data constraints, none of the estimation strategies can solve all estimation issues. As a result, in the absence of a dominating specification, we use a set of strategies to establish an estimator-robust direction of parenting skills and time input impacts on a child's skill development. To do so, we proceed with the following:

1. Estimate the impact of different parenting dimensions on skill development
2. Identify if there is a quantity-quality trade-off between time investments and parenting style
3. Estimate if time investments have differential impacts depending on parenting style

4. Estimate if parenting style impacts skills in a non-linear way

In the next paragraphs we detail on each step of the estimation strategy. For the first step, we are going to use the set of six econometric models to estimate the impact of parenting dimensions for which we are going to discuss identifying assumptions. Our main specification takes the following form:

$$Y_{ia} = \alpha_a + PS'_{ia}\delta_a + TI'_{ia}\gamma_a + R'_{ia}\rho_a + \epsilon_{ia} \tag{2.2}$$

where current skills Y_{ia} depend on current parenting style PS_{ia} and time investments TI_{ia}. Further, R_{ia} describes all other relevant inputs, not including current parenting style and time investments. This term could entail initial endowment, past investments, and other observable characteristics.

In terms of econometric models, we start with the most simple model, the contemporaneous linear model (OLS):

$$Y_{ia} = \alpha_a + PS'_{ia}\delta_a + TI'_{ia}\gamma_a + X'_{ia}\beta_a + \epsilon_{ia} \tag{2.3}$$

where current skills only depend on current parenting style PS_{ia}, time investments TI_{ia}, and characteristics X_{ia}. Household characteristics in our base specification are the age of the primary caregiver, the number of siblings, the log of family income and dummies for college education of the primary caregiver, the presence of both biological parents at home, the gender of the study child, the indigenous status of the child, the study child's cohort, and the day of data collection. This specification is only unbiased if production inputs are constant over time. Then, the current values can summarize the whole history of production inputs. Further, the current inputs are uncorrelated with, for instance, the permanent unobserved ability of parents, or temporary shocks, which would be captured by the error term ϵ_{ia}. Additionally, current characteristics X_{ia} need to proxy well for the innate ability Y_{i0} in Equation 2.1.

To control for innate ability and past investments we add lagged skill measures to the estimation, employing a Value Added model (VA) as a second model:

$$Y_{ia} = \alpha_a + PS'_{ia}\delta_a + \lambda_a + TI'_{ia}\gamma_a + X'_{ia}\beta_a + Y_{ia-1} + \epsilon_{ia} \tag{2.4}$$

This model is based on the assumption that the past period's skill outcomes Y_{ia-1} capture the impact of past investments and innate ability with rate λ_a. Additionally, to be unbiased, all current investments that respond to past skills must be part of the estimation. Further, if Y_{ia-1} is measured with a measurement error correlated with the one of Y_{ia}, this might exacerbate measurement error bias.

Another way to control for unobserved ability is using fixed effects (FE), our third model:

$$Y_{ia} = \alpha_i + PS'_{ia}\delta + TI'_{ia}\gamma + X'_{ia}\beta + \epsilon_{ia} \tag{2.5}$$

where α_i is a child-fixed effect. This approach assumes that the fixed effect captures the child's innate ability and other time-invariant influences as parents' ability. However, this estimation strategy only leads to unbiased estimates if the impact of these time-invariant factors is constant across ages. Additionally, past investments do not influence current skills after controlling for innate ability or do not correlate with current investments like parenting style or innate ability. Another assumption is strict exogeneity: past, current, and future inputs are not correlated with past, current, and future errors.

To allow for influence of past investments on current skills, we extend the FE framework by controlling for lagged skills. This gives us the fourth model (AB):

$$Y_{ia} = \alpha_i + PS'_{ia}\delta + TI'_{ia}\gamma + X'_{ia}\beta + \lambda Y_{ia-1} + \epsilon_{ia} \tag{2.6}$$

We estimate this model using a GMM estimation that uses all available exogenous variation in estimation proposed in the seminal paper of Arellano and Bond (1991). This allows us to relax the strict exogeneity assumption required for the fixed effects model. The Arellano-Bond estimator introduces a weaker assumption, the orthogonality condition, which states that the lagged skill measure is uncorrelated with the error term after controlling for the lagged production inputs. This means that the lagged skills can be used as an instrumental variable to address the endogeneity issue caused by the presence of lagged skills in the model.

Another approach is to control for all past inputs, which gives us the fifth estimation strategy, the cumulative model (CU):

$$Y_{ia} = \alpha_a + \sum_{t=0}^{a} PS'_{it}\delta_{at} + \sum_{t=0}^{a} TI'_{it}\gamma_{at} + X'_{ia}\beta_a + \epsilon_{ia} \tag{2.7}$$

This specification controls for all available past inputs, however, not for innate ability. Therefore, the assumption is that either innate ability is uncorrelated with past and current inputs or captured well by past investments. To control for innate ability in the sixth and last specification, we again add lagged skill outcomes. This gives us the cumulative model with lagged inputs and skills (CV):

$$Y_{ia} = \alpha_a + \sum_{t=0}^{a} PS'_{it}\delta_{at} + \sum_{t=0}^{a} TI'_{it}\gamma_{at} + X'_{ia}\beta_a + \lambda_a Y_{ia-1} + \epsilon_{ia} \tag{2.8}$$

This model relies on the assumption, that all investments are controlled for as otherwise they might be in the error term and bias coefficients of interest. Additionally, measurement error in skills, might effect results and the assumption is that innate ability influences skills at rate λ_a.

After running these models for the main specification in Equation 2.2, we proceed by the following specifications with the same subset of econometric models. Firstly, that is the main specification but only with time investments, to see if they alone can explain skill development and how magnitudes change with adding parenting style to the equation to see if results are robust:

$$Y_{ia} = \alpha_a + TI'_{ia}\gamma_a + R'_{ia}\rho_a + \epsilon_{ia} \tag{2.9}$$

Secondly, we run a specification including interaction terms between current time investments and parenting style, to see if time investments have differential impacts depending on parenting style and if there is a quantity-quality trade-off:

$$Y_{ia} = \alpha_a + PS'_{ia}\delta_a \times TI'_{ia}\gamma_a + +R'_{ia}\rho_a + \epsilon_{ia} \tag{2.10}$$

Using interactions between five initially defined parenting styles and four time investments would result in many interaction terms (20 in total). It can lead to statistical inefficiency and potential collinearity issues. Further, the high dimensionality can make the interpretation and estimation of the model more complex and challenging. Instead, we aggregate the parenting styles into two broader dimensions. Section 2.2.1 describes the construction of these parenting styles, and Table B.3 displays the result of the factor analysis used to derive them. Then, we use aggregated parenting styles to create interactions with time investments. This approach simplifies the model by reducing the number of interaction terms to only eight (2 parenting dimensions multiplied by 4 time investments). It reduces the risk of multicollinearity and makes the estimation more manageable.

Last but not least, we test for non-linearities in the impact of parenting style dimensions on skills:

$$Y_{ia} = \alpha_a + PS'_{q2,a}\delta_{q2,a} + PS'_{q3,a}\delta_{q3,a} + PS'_{q4,a}\delta_{q4,a} + TI'_{ia}\gamma_a + R'_{ia}\rho_a + \epsilon_{ia} \tag{2.11}$$

To do so, we devide the sample of parents into quartiles ($PS_{q1,a}, PS_{q2,a}, PS_{q3,a}, PS_{q4,a}$) to determine if extreme forms of parenting have a particularly strong impact on skill development. We include the 2nd to 4th quartile dummy in the estimation and keep the 1st quartile as base category.

2.4. Results and Discussion

We now discuss the estimation results obtained following the strategy discussed in Section 2.3. We start with the results for non-cognitive skills, followed by those for cognitive skills, and briefly discuss the outcomes of the conducted robustness checks. For both types of skills, we first present the results of the main specification in Equation 2.2, and the implications of estimating this equation using different econometric models (see Equations 2.3 - 2.6). We also compare these results to those obtained from estimating the skill development process only with time investments as inputs, without considering parenting style (see Equation 2.9). Second, we highlight the outcomes of interacting parenting styles with parenting time, as described in Equation 2.10. Thirdly, we discuss the results of testing for non-linearities (Equation 2.11).

Non-cognitive skills

We start by describing results for non-cognitive skills using the main specification in Equation 2.2. Tables 2.3 and B.10 - B.12 summarize the different econometric models for each age group at which the skill development process is estimated. To facilitate interpretation, we will use a specific age group as the benchmark and then compare the outcomes to other age groups. Specifically, we examplary choose to present mainly estimation results for the age group 8-9, as shown in Table 2.3 (arbitrarily chosen as the first wave with data available for all specifications). Later we will relate them to other age groups.

Table 2.3 shows that hostile parenting is consistently negative and significant at conventional levels across different econometric models. Similarly, parental reasoning and inconsistent parenting negatively influence skill development (parental reasoning becomes insignificant in the Arellano-Bond model). The magnitudes vary, with hostile parenting having the biggest impact on non-cognitive skills, followed by inconsistency and a lower impact on reasoning. Increasing hostility by one standard deviation (SD) leads to a decrease in non-cognitive skills by 0.12-0.35 SD, depending on the model employed. Inconsistency, increased in the same magnitude, decreases skills by 0.04-0.15 SD, and reasoning by 0.01-0.08 SD. Parental warmth has a small positive impact across models, varying in significance. Attempted consistency seems not to affect skill outcomes, indicating that what matters is the enforced consistency, summarized under the factor named inconsistent parenting. Regarding the self-productivity of skills, we find evidence for it for non-cognitive skills. We find estimates for the impact of last periods test outcome to be between 0.25 and 0.64 SD, indicating that high levels of non-cognitive skills in earlier ages persist to some extend independent of current investments and parenting style. These results are consistent across age-groups.

The literature suggests that children of authoritative parents tend to have higher non-cognitive skills (Spera (2005), Luyckx et al. (2011), Delvecchio et al. (2020), McWhirter et al. (2023)). The authoritative parenting style is characterized by high warmth, reasoning, and consistency, and low hostility. Our results support these findings, with warmth having a tentative positive effect and hostility and inconsistency having negative effects. Moreover, our analysis enables us to identify which dimensions of the authoritative parenting style are associated with higher non-cognitive skills. Our findings suggest that low hostility is the primary driver of the positive impact of this style, followed by high consistency. In contrast, warmth appears to have a limited role, while reasoning may have a negative effect.

The literature has also found negative associations between non-cognitive skills and authoritarian and neglectful parenting styles (Fiorini and Keane (2014), Le Forner (2021), Spera (2005), Fletcher et al. (2008), Heberle et al. (2015), McWhirter et al. (2023)). Our results suggest that this negative association might be due to high hostility levels, offsetting the positive impact of consistency for authoritarian parenting. For neglectful parenting, low consistency and warmth might contribute to the negative association with skills.

Permissive parenting is associated with more externalizing problems and antisocial

Table 2.3: Estimated parameters of production function for non-cognitive skills at age 8-9

	OLS	VA	FE	AB	CU	CV
Parental warmth	0.065***	0.020	0.048**	0.011	0.050*	0.042**
	(0.022)	(0.017)	(0.020)	(0.017)	(0.027)	(0.020)
Parental reasoning	-0.084***	-0.043***	-0.025*	-0.013	-0.056***	-0.031*
	(0.019)	(0.014)	(0.015)	(0.015)	(0.021)	(0.016)
Hostile parenting	-0.352***	-0.163***	-0.162***	-0.117***	-0.295***	-0.188***
	(0.023)	(0.018)	(0.016)	(0.017)	(0.026)	(0.020)
Inconsistent parenting	-0.153***	-0.060***	-0.061***	-0.044**	-0.108***	-0.074***
	(0.023)	(0.018)	(0.018)	(0.018)	(0.028)	(0.021)
Attempted consistency	0.003	-0.011	-0.017	-0.010	-0.010	-0.007
	(0.018)	(0.015)	(0.015)	(0.014)	(0.021)	(0.016)
Educational time parents	-0.002	-0.000	0.001	0.003	-0.003	0.001
	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)	(0.003)
Educational time others	0.006	0.000	-0.013	-0.012	0.010	0.001
	(0.014)	(0.012)	(0.010)	(0.010)	(0.015)	(0.012)
Care time parents	-0.002	-0.001	-0.001	-0.000	-0.001	-0.000
	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
Care time others	0.004	-0.001	-0.000	-0.003	0.004	-0.002
	(0.007)	(0.004)	(0.004)	(0.004)	(0.008)	(0.005)
Lagged test outcome		0.637***		0.253***		0.635***
		(0.020)		(0.019)		(0.020)
Observations	2,780	2,667	6,599	6,463	2,419	2,417
Lagged dependent	NO	YES	NO	YES	NO	YES
Lagged inputs	NO	NO	NO	NO	YES	YES
Fixed effects	NO	NO	YES	YES	NO	NO

Notes: Each specification includes controls as the age of the primary caregiver, the number of siblings, the log of family income, and dummies for the college education of the primary caregiver, the presence of both biological parents at home, the gender of the study child, the indigenous status of the study child, the study child's cohort, and the day of data collection.

behavior (see McWhirter et al. (2023) for an overview). Regarding our findings with regard to non-cognitive skills, in the case of permissive parenting, the negative impact of inconsistency might be offset by low levels of hostility and higher warmth, depending on the magnitude of these dimensions. In general, it seems promising to target parental behaviors that lead to hostility and inconsistency in parenting training to increase the effectiveness of these interventions.

Figure 2.3 shows how different parenting styles affect the non-cognitive skills of children in different age groups. The graph shows the predicted values of non-cognitive skills at the 25th, 50th, and 75th percentiles of parenting styles, using coefficients from a fixed effects model with lagged values (Arellano-Bond). Generally the skill gap is the widest across at all ages within the hostile parenting group. While analysing the different age groups, the gap in skills between children with hostile parents (75th percentile) and those with non-hostile parents (25th percentile) is smallest among the youngest age group and increases slightly over time. At the age of 14-15, the difference in non-cognitive skills between children with hostile parents in the 25th percentile and those in the 75th percentile is about 0.2 standard deviations. The impact of parental warmth, reasoning, and inconsistent parenting on non-cognitive skills also increases with age. However, the difference in non-cognitive skills between children in the 25th and 75th percentiles of these parenting styles is much smaller than in the case of hostile parenting, and in some cases, it is insignificant. Lastly, attempted consistency in parenting does not appear to affect non-cognitive skills in every age group. To sum up, these graphs illustrate that, in particular, hostile parenting has a negative impact on non-cognitive skills, especially at later ages.

Table B.13 presents our estimation results for Equation 2.9 using only time investments as inputs for the age group 8-9 to see if estimates of time investments are robust to including parenting style. The coefficients for all time investments remain insignificant, indicating no significant differences in the impact of different types of time spent with the child compared to other time investments. Additionally, the coefficients do not vary much in magnitude, indicating they are robust to including parenting dimensions in the specification. These findings are consistent across age groups (see Tables B.14 - B.16) and align with previous studies by Fiorini and Keane (2014) and Le Forner (2021), who also did not find significant impacts of time investments on non-cognitive skills on younger children (age 4-11). We extend this analysis to older children (middle childhood/teenage years), and our results align with the literature. Del Boca et al. (2014) find a decreasing productivity of active time spent with children by age in the UK, further supporting our findings.

Next, we investigate the interaction of time investments with parenting style. Table B.21 displays the estimates of this specification (Equation 2.10) for age group 8-9 (for estimates for other groups, see Tables B.22 - B.24). Note that here we use the parenting styles to summarize parenting dimensions. We also run the main specification, Equation 2.2, using summarizing styles instead of the parenting dimensions. Using the summarizing

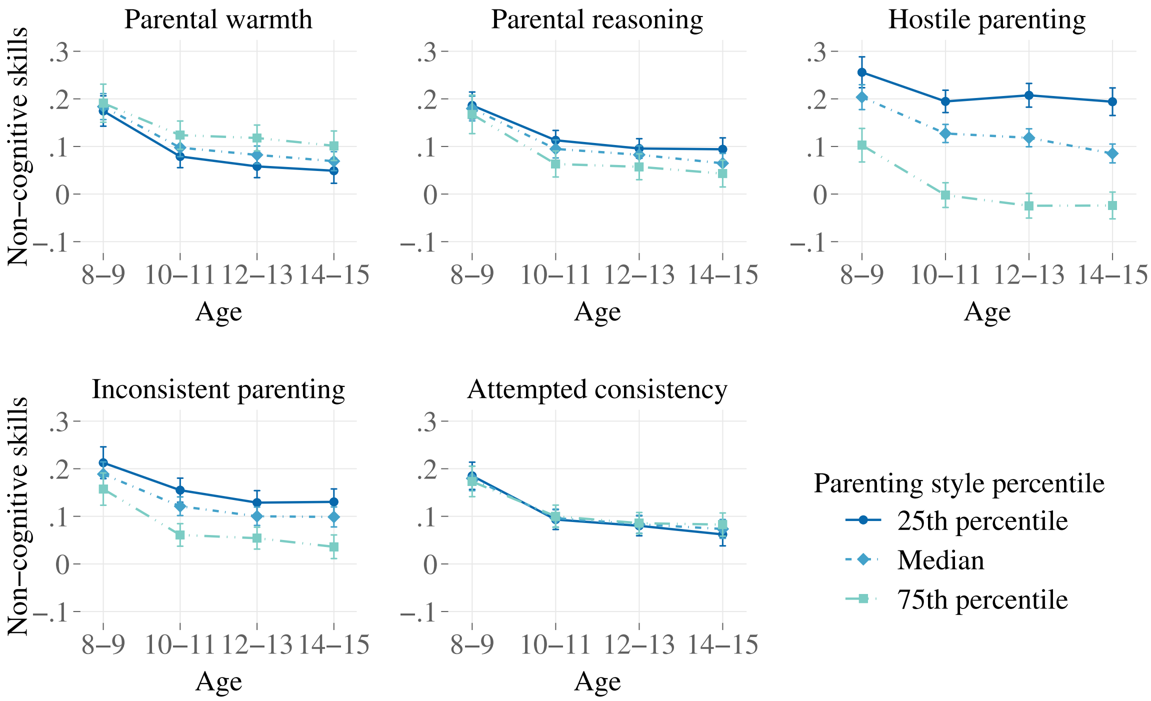


Figure 2.3: Change in the impact of parenting dimensions over age

Notes: The figure presents predicted values of standardized non-cognitive skills for the 25th, 50th, and 75th percentile of the different parenting styles. The rest of the production inputs are at the sample mean. The range bars correspond to a 95% confidence interval for the point prediction.

styles leads to similar conclusions as the styles originating from the single factor analysis. Therefore, the summarizing scores seem to reflect the findings for the parenting dimensions well (see Tables B.17 - B.20 for details). Turning to the interactions, we do not find any additional effect of parenting style with increasing time investments at any age. This is the case for an empathetic style (capturing high reasoning, warmth, and attempted consistency) and a harsh style (capturing hostility and inconsistency). This shows that including parenting style as an additional input independent of time investments would better mimic the skill development process.

To test for non-linearities, we estimate Equation 2.11. Using dummies for each quartile of parenting style, we can explore if extreme values disproportionally impact skills. Figure 2.4 presents estimated coefficients of dummies for belonging to the quartiles of different parenting dimensions (with the 1st quartile as a baseline category) for four age groups (for estimates, see Table B.25 - B.28). It allows us to analyze the potentially non-linear relationship between parenting dimensions and non-cognitive skills. The results suggest a non-linear relationship between some parenting styles and non-cognitive skills. In general, the non-linearity is stronger for older age groups. However, the nature of this relationship varies by the dimension of the parenting style. Regarding parental warmth, dummies for the quartiles are mostly insignificant, which aligns with findings from the estimation with the linear specification. Moving from one quartile to another has a decreasingly negative impact on non-cognitive skills in the case of parenting reasoning. In contrast, for hostile and inconsistent parenting, moving to a higher quartile has an increasingly negative impact on non-cognitive skills, indicating that parents with high hostility and inconsistency have an especially detrimental impact on their child's non-cognitive skills. In the case of attempted consistency, most of the coefficients are insignificant, which confirms the findings drawn from the linear specification.

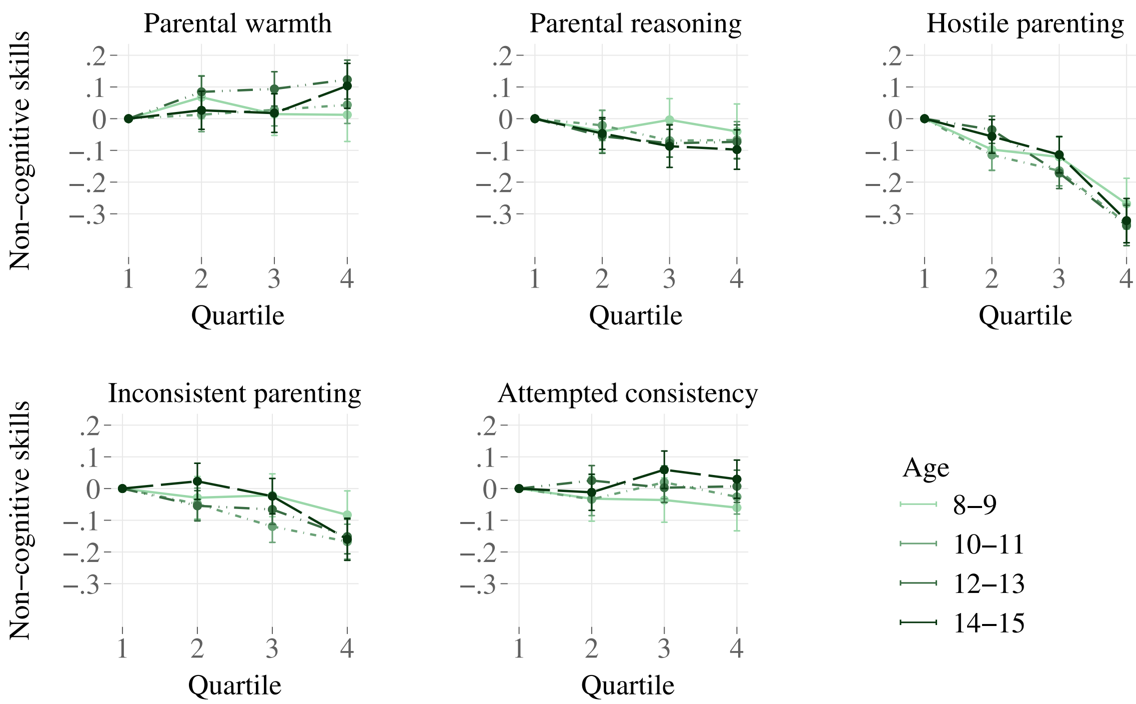


Figure 2.4: Non-linear impact of parenting style on non-cognitive skills

Notes: The figure presents estimated coefficients of dummies for belonging to the 2nd, 3rd, and 4th quartile of parenting styles (the 1st quartile serves as baseline category) for four age groups: 8-9, 10-11, 12-13, and 14-15. The range bars correspond to a 95% confidence interval for the estimated coefficients.

Cognitive skills

Tables 2.4 and 2.5 present the estimation results for the cognitive skills production function at age group 8-9 (see Table B.29 for estimation results for age group 10-11). Tables 2.4 display the results for MRT scores, while Table 2.4 shows the results for PPVT scores. Most of the coefficients for parenting dimensions are insignificant, except for inconsistency and warmth in some specifications. The magnitudes of the coefficients are small, and the standard errors indicate that the impact is likely to be zero rather than noisy estimates. Inconsistency and warmth have negative coefficients, suggesting they are associated with a decrease in cognitive skills. However, overall, parenting style does not seem to substantially impact cognitive skill development. In contrast, educational time spent with parents positively affects skills in most econometric models, particularly for the younger age group (8-9). The coefficients are not significant in the estimations using the fixed effect and Arellano-Bond approach. It might suggest that the effect of educational time with parents disappears when one controls for the child's innate ability. Excluding parenting style from the estimation does not significantly alter the coefficients for time investments (see Tables B.30 - B.32), and results are similar for all age groups. Regarding the self-productivity of skills, we also find evidence for it for cognitive skills. Estimates for the impact of last periods test outcome are between 0.18 and 0.46 SD for the MRT and 0.15 and 0.49 SD for the PPTV. These values are slightly lower than in the case of non-cognitive skills, suggesting a stronger persistence of non-cognitive skills.

These results align with Fiorini and Keane (2014) and Le Forner (2021), who do not find evidence for the impact of an authoritarian or warm parenting style on cognitive skills. Similar to their analysis, we find evidence for the impact of educational time spent with parents on cognitive skills. In general, our evidence of the impact of different types of time investments is weaker than the one found by Fiorini and Keane (2014) and Le Forner (2021). However, we look partly at older cohorts, where time investments matter less (see Del Boca et al. (2014)). Additionally, time investment measures vary across ages due to the survey collection method, which could drive these results to some extent.

We further test for the effect of parenting style on additional time spent with the child. Results for the interaction of parenting styles with time investments are displayed in Tables B.33 - B.35. We do not find consistent effects of the interaction between parenting style and time investments. Hence, for the skill formation of cognitive skills, skill production functions without parenting style are likely to capture the skill process well in contrast to non-cognitive skills. Similar conclusions hold for non-linearity. Tables B.36 - B.38 show the estimation results for testing for non-linearity in the impact of parenting styles. We do not see strong evidence for non-linearity.

These results indicate that parenting training targeting parenting style might be particularly effective in increasing non-cognitive skills but not cognitive ones. Depending on which improvements policymakers aim for, different intervention designs are needed. However, it is important to keep in mind that severe behavioral problems can impact grade progression and school outcomes which in return might affect cognitive skill development in the long run and that there are increasing returns from non-cognitive skills for wages later in life (see Carneiro et al. (2007), Deming (2017) and Edin et al. (2022)).

Robustness checks

In this section, we examine the robustness of our findings on the relationship between parenting style and non-cognitive and cognitive skills. We analyse the sensitivity of our results to various econometric specifications. By conducting these additional analyses,

Table 2.4: Estimated parameters of production function for cognitive skills (MRT) at age 8-9

	OLS	VA	FE	AB	CU	CV
Parental warmth	-0.040	-0.036	-0.034	-0.017	-0.033	-0.035
	(0.027)	(0.025)	(0.024)	(0.033)	(0.034)	(0.031)
Parental reasoning	-0.002	-0.003	-0.005	-0.007	-0.017	-0.016
	(0.024)	(0.021)	(0.021)	(0.028)	(0.027)	(0.025)
Hostile parenting	0.004	0.014	0.014	-0.002	0.008	0.018
	(0.026)	(0.023)	(0.023)	(0.030)	(0.032)	(0.028)
Inconsistent parenting	-0.068***	-0.052**	-0.027	0.003	-0.037	-0.027
	(0.026)	(0.023)	(0.024)	(0.031)	(0.031)	(0.026)
Attempted consistency	0.002	-0.015	-0.032*	-0.029	-0.032	-0.050**
	(0.022)	(0.019)	(0.019)	(0.025)	(0.027)	(0.023)
Educational time parents	0.014***	0.007*	-0.000	0.003	0.014***	0.006
	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.004)
Educational time others	0.017	0.013	0.012	0.007	0.001	0.001
	(0.017)	(0.016)	(0.015)	(0.016)	(0.017)	(0.015)
Care time parents	-0.000	0.000	0.002	0.003	-0.001	-0.000
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Care time others	0.005	0.003	0.002	0.009	0.012*	0.006
	(0.006)	(0.006)	(0.005)	(0.006)	(0.006)	(0.006)
Lagged test outcome		0.457***		0.179***		0.455***
		(0.019)		(0.039)		(0.021)
Observations	2,753	2,690	7,428	2,504	2,399	2,392
Lagged dependent	NO	YES	NO	YES	NO	YES
Lagged inputs	NO	NO	NO	NO	YES	YES
Fixed effects	NO	NO	YES	YES	NO	NO

Notes: Cognitive skills are measured using Matrix Reasoning Test outcomes standardized by age group. Each specification includes controls as the age of the primary caregiver, the number of siblings, the log of family income, and dummies for the college education of the primary caregiver, the presence of both biological parents at home, the gender of the study child, the indigenous status of the study child, the study child's cohort, and the day of data collection.

Table 2.5: Estimated parameters of production function for cognitive skills (PPVT) at age 8-9

	OLS	VA	FE	AB	CU	CV
Parental warmth	-0.050**	-0.048**	-0.070***	-0.055*	-0.030	-0.037
	(0.024)	(0.021)	(0.026)	(0.028)	(0.031)	(0.026)
Parental reasoning	0.019	0.016	-0.001	-0.002	-0.003	-0.017
	(0.024)	(0.020)	(0.025)	(0.026)	(0.027)	(0.023)
Hostile parenting	-0.001	-0.014	-0.039	-0.035	0.020	0.002
	(0.023)	(0.020)	(0.025)	(0.027)	(0.028)	(0.025)
Inconsistent parenting	-0.091***	-0.035*	0.014	0.020	-0.053*	-0.025
	(0.023)	(0.021)	(0.026)	(0.026)	(0.028)	(0.026)
Attempted consistency	0.025	-0.004	-0.011	0.004	0.004	0.001
	(0.021)	(0.018)	(0.021)	(0.023)	(0.024)	(0.021)
Educational time parents	0.017***	0.010***	0.001	-0.004	0.015***	0.008**
	(0.004)	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)
Educational time others	0.013	0.001	-0.016	0.002	0.003	-0.007
	(0.015)	(0.015)	(0.025)	(0.025)	(0.016)	(0.016)
Care time parents	0.002	0.000	-0.001	-0.002	-0.000	-0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Care time others	-0.001	-0.003	-0.004	0.003	0.006	0.001
	(0.005)	(0.005)	(0.010)	(0.011)	(0.006)	(0.006)
Lagged test outcome		0.494***		0.150***		0.488***
		(0.019)		(0.044)		(0.020)
Observations	2,755	2,633	3,437	2,156	2,401	2,343
Lagged dependent	NO	YES	NO	YES	NO	YES
Lagged inputs	NO	NO	NO	NO	YES	YES
Fixed effects	NO	NO	YES	YES	NO	NO

Notes: Cognitive skills are measured using Peabody Picture Vocabulary Test (PPVT) outcomes standardized by age group. Each specification includes controls as the age of the primary caregiver, the number of siblings, the log of family income, and dummies for the college education of the primary caregiver, the presence of both biological parents at home, the gender of the study child, the indigenous status of the study child, the study child's cohort, and the day of data collection.

we aim to ensure that our conclusions are robust to model specification.

First, we examine how our results are affected by the way we define parenting styles. In the main specification, we conducted a factor analysis on selected subgroups of questions related to parental behavior, resulting in five different dimensions of parenting styles. Alternatively, we can pool all the questions together and obtain four factors in a joint factor analysis. Based on the reported loadings (see Tables B.4 - B.9), we labeled the factors as warm style, reasoning style, hostile and inconsistent style, and consistent style. Depending on wave, the described factors load differently, so we assign the factors produced to the fitting variable (see table notes for details). We then run our main specification with the jointly estimated factors.

Tables B.39 - B.42 present the estimated parameters of the production function with the jointly estimated factors for non-cognitive skills. On average, warm style and consistent style have a positive influence on children's non-cognitive skills. However, the impact of the latter one is insignificant in the specifications with fixed effects and becomes negative for the age group 14-15. In contrast, reasoning style and hostile and inconsistent style have a negative impact on non-cognitive skills. These findings are consistent with the previous results. The main difference is that the jointly estimated warm style has a consistently significant impact, which is not the case in the main specification. The reason for this might be that in the joint factor case, warmth indicates among others how much parents praise their child. In the case of the main specification with parenting dimensions, praise is classified under hostility given the survey module. This indicates that praising positively affects the development of non-cognitive skills. Additionally, in the jointly estimated model, the consistent style has an impact and is significant in some model specifications. This could additionally support the conclusion that implemented consistency matters and not attempted one, which is mainly insignificant in the main specification.

Tables B.47 - B.49 present the estimated production function parameters with jointly estimated factors for cognitive skills. The effect of parenting style is, on average, small and mostly insignificant. This confirms previous findings that parenting style does not strongly impact a child's cognitive skill development.

As a second robustness check, we include an extended set of controls in the main model specification. The additional controls consist of dummies for urban areas and Australian states, aimed at capturing the effect of geographical factors on the development of non-cognitive and cognitive skills. They may also account for differences in school quality and public services between rural and urban areas and across states. The second control group consists of the Socio-Economic Indexes for Areas (SEIFA), which measure four aspects of socio-economic advantage and disadvantage. These indexes are constructed based on Australian census data. They are created for each statistical area, allowing us to control the economic situation of the local region in which the child is growing up. Finally, we also include controls for birth weight and early birth as proxies for innate ability.

The addition of a new set of controls only marginally changes the estimated coefficients associated with parenting styles and time investments in the production function of non-cognitive skills (see Tables B.43 - B.46). Similar conclusions can be drawn from the analysis of the same coefficients in the production function of cognitive skills (see Tables B.50 - B.52). This indicates that the omitted controls in the main specification are uncorrelated with the parenting style inputs and do not represent a source of endogeneity bias.

Despite these robustness checks, certain limitations of our approach remain. Firstly given that we use time-invariant fixed effects, we do not control for time-varying selection. Therefore, estimates could be biased if, for example, financial shocks influence skills directly and indirectly via increasing parenting hostility due to stress. Another limitation is the measurement of parenting skills and non-cognitive skills. Firstly, there might be measurement errors. Measurement error might be tackled using the latent factor modeling approach of Cunha and Heckman (2008) and Cunha et al. (2010) in future extensions of this work. Further, we use measures of cognitive skills and parenting style, which are self-reports by parents. García-Miralles and Gensowski (2023) point out that parental health shocks influencing their children's health might change parents' perception of their child's behavior at the same time. A similar endogeneity might occur in our case. Hostile parents could perceive their children as more hyperactive than others because they might value obedience higher than permissive parents and notice it more. One could extend this work in the future by using teacher reports to validate parents' reporting in a robustness check. Further, our measures of time investments are not consistent over waves as the method of reporting changes, as well as the responding person. This could drive the results of the interaction of time investments with skills to some extent. If the varying measures do not capture time investments well, measurement error might bias results.

2.5. Summary and Concluding Remarks

In this paper, we investigate the impact of different parenting dimensions on cognitive and non-cognitive skill development. Additionally, we provide empirical facts showing the association between income and parenting dimensions. To do so, we use the Longitudinal Survey of Australian Children to estimate the impact of parental warmth, reasoning, hostility, and consistency on skills. We exploit the panel structure of the data and the availability of rich demographic and investment variables to control for potential endogeneity issues. These include controlling for unobservable time-invariant characteristics, past investments, and skill outcomes. Doing so, we enrich the existing literature by providing a structured analysis of the impact of parenting dimensions on skills testing for interactions with time investments and non-linearities.

We find that non-cognitive skills decrease with higher parenting hostility and inconsistency and to a lesser extent with higher reasoning. Parenting warmth positively influences non-cognitive skills, however, with low magnitude and not consistently significant across the employed econometric models. We show that the positive association between authoritative parenting and skill development found in the literature seems to be driven by low levels of hostility and inconsistency. Parenting warmth and reasoning play a limited role. In contrast, for authoritarian parents, the higher level of hostility seems to offset the positive effect of consistency. We find hostility impacts skill development in higher magnitudes than inconsistency and that the impact increases with age. As parents from the bottom of the income distribution tend to have higher levels of hostile and inconsistent parenting, these factors might be an additional driver for the skill gap between children from different socioeconomic backgrounds (at least for non-cognitive skills).

Regarding how to model parenting style in skill formation, we do not find evidence for parenting styles influencing the impact of time investments, e.g., increasing their productivity. Therefore, parenting style should be modeled as additional investment input in skill production functions. Further, hostile parenting is the only parenting dimension displaying a strong non-linear relationship in impact on non-cognitive skills. Hence, our analysis indicates that linear modeling of the impact of parenting styles captures the skill formation process well (except for hostility). We do not find consistent evidence for the influence of parenting styles on cognitive skills. For the development of cognitive skills, other factors seem to be more important, like as the literature suggest, time investments or monetary investments (see Del Boca et al. (2014)). This highlights the importance of modeling non-cognitive and cognitive skill development with different functional form assumptions and inputs.

These results indicate that for non-cognitive skill development, it is particularly important to have parents with a low level of hostility and inconsistency. This finding is informative for the design of child development policies. For instance, targeting these two parenting behaviors might be particularly efficient regarding parenting training. Given the non-linearity of the impact of hostility, it might be particularly important to target parents who display a high level of hostility or are likely to do so. Given that we find the impact of hostility to increase with age, targeting adolescence seems important. Nonetheless, given that skills are self-productive, starting at earlier ages could be beneficial, in particular as we find non-cognitive skills to display more persistence than cognitive skills. More research is needed to determine the trade-off between periods. Hostile and inconsistent parenting is often associated with increased stress levels in parents. Therefore, another promising approach might be to combine parenting training with stress management training to maximize the impact. Nonetheless, more research is needed on the amenability of these behaviors to determine the efficiency of this approach, and our results indicate that focusing, in particular, on hostility and inconsistency in doing so is promising.

3. ACROSS-DISTRICT MARRIAGE MIGRATION IN INDIA

3.1. Introduction

Within-household inequality is responsible for 30% of total inequality in India (Klasen and Lahoti, 2021). Allocation of resources within the household is, on average, unfavorable for women partners which results in their worse health conditions or higher mortality risk (Calvi, 2020). This inequality in access to household resources has economic, cultural, and social origins. One of the contributing factors could be associated with the migration patterns. While most men migrate for work, 65% of women migrate for marriage[1], leaving the parental home to join the groom's family. Of these, a fifth (21%) have migrated across districts (but within the same state), which requires crossing cultural borders. The practice of importing brides from another district might be driven by the shortage of eligible brides in certain areas, as a result of skewed sex ratios. As a consequence, there is a considerable regional variation in the number of marriage migrants from 8% to 40%.

While there is some evidence about across-district marriage migration and the causes of it (Fulford, 2013, 2015, Kaur, 2008, Rosenzweig and Stark, 1989), little is known about its consequences for the within-household inequality. On the one hand, women moving to another district might suffer from increasing distance from the parental home. They leave behind their support structures and are exposed to discrimination due to cultural differences. On the other hand, women usually move to regions with more skewed sex ratios, which should increase their bargaining power, in theory. As it is not clear how these two mechanisms play out, in this paper, we investigate the consequences of across-district marriage migration in India for within-household inequality. Specifically, we ask: does the across-district marriage migration help or hurt women? In particular, does it reduce or increase female bargaining power? As a result, does it enlarge the overall level of inequality?

To answer these questions, we first investigate the correlation between across-district migration and geographical variations in the sex ratio, shedding light on the relationship between these factors. Using logistic regression, we explore how individual and spouse's household characteristics influence the probability of women migrating to another district. This analysis provides valuable insights into the determinants of migration decisions in the Indian marriage market. Furthermore, we build a static marriage market model with transferable utility in the spirit of Choo and Siow (2006). We extend the standard framework by incorporating marriage migrants, which allows for a comprehensive examination of their effects on women's and men's marriage surpluses. Finally, we develop a theoretical model of the collective household that offers a foundation for future detailed analyses of changes in bargaining power due to marriage migration.

[1]Source: Census of India 2001

Conducting empirical analyzes requires overcoming several challenges. First, marriage migration decisions are endogenous, influenced by the marriage market dynamics in neighboring districts. Simultaneously, migrating women impact these market conditions. Second, we usually do not observe the marriage market participation. Instead, we only have access to marriage market outcomes, such as marriage patterns. This makes establishing an equilibrium in the marriage market more challenging due to the lack of complete knowledge of who actively participates in the marriage market. Consequently, it requires additional assumptions to compensate for the limited information.

In addition to methodological challenges, we also need to address data limitations. Answering these research questions requires data that captures the history of migration decisions and variables that allow the identification of bargaining power within households. However, no single dataset contains all this information for India. It requires a multi-method approach that allows to combine information from different datasets.

For this reason, we use two data sources. The first is the Census of India, conducted every ten years. This dataset allows us to calculate the sex ratio for 5 year age group intervals at the district level. The second data source is the National Sample Survey (NSS), a household survey conducted by the Ministry of Statistics and Programme Implementation of India. This survey contains information about basic demographic characteristics of all household members as well as their entire migration history. This feature allows us to identify women who migrated for marriage and the type of migration. We use the 2001 year for the Census and 2006/07 for the NSS.

We tackle the challenges of studying marriage migration in India through a multi-method approach. The first step involves estimating logistic model to understand the determinants of migration decisions. Using a reduced form approach allows us to examine the characteristics of women who choose to migrate across districts and the types of households that attract them and thus, we are able to provide insights into the factors influencing marriage migration patterns in the Indian context.

Further, we employ two models to study the effect of marriage migration on intra-household bargaining power. We start following approach of Choo and Siow (2006) and construct a static marriage market model with transferable utility. This is a standard approach in the literature to estimate marital gains. However, it requires data on all participants in the marriage market in order to identify the marriage surpluses. Since these data are not available in our case, we make assumptions regarding women who might participate in the district's local marriage market but come from other districts of the same state. By analyzing these surpluses, we shed light on the effects of marriage migration on the division of marriage surplus between men and women. The static model used in this study may not fully capture all the relevant aspects of the impact of marriage migration on women's position within the household, primarily due to data limitations. Consequently, we propose a second model that employs a collective household approach to estimate the bargaining power at the individual level. This model provides the framework to test three fundamental hypotheses: the correlation between bargaining power and the probability of marriage migration, the relationship between bargaining power and the size of the dowry, and the association between dowry and the probability of marriage migration. Testing these hypotheses allows us to identify the crucial trade-off behind the marriage migration: lower bargaining power for a price of lower dowry.

Our descriptive analyses provides two important facts: (a) a positive correlation between across-district marriage migration and the sex ratio in the district: as the sex ratio worsens marriage migration increases. (b) a positive correlation between across-district marriage migration and the state sex ratio. The states that have worse sex ratio also have more migration between their districts for marriage purposes. Next, using logistic regressions we show that marriage migrants are more likely to have at least primary education and to move to a household where the head also has at least primary education (Rao and Finnoff, 2015). Comparing amongst socioeconomic and religious groups, we find members of scheduled tribes to be less likely to migrate for marriage purposes to another district, relative to other groups. Finally, the probability that a woman in the household migrated for marriage increases with household per capita consumption expenditures, suggesting that women move to more wealthy households. From this analyses, it is unclear whether there is a positive or negative impact of marriage migration on woman position in the household.

Therefore, using a modified Choo and Siow (2006) model, we uncover men and women preferences for within and across-district migration and focus on the division of the marriage surplus that might be linked to the within-household inequality. We find that: (a) states with higher marriage migration also have higher across-district surplus; (b) with the increase in marriage migration, within-district marital surplus declines, while across-district marriage surplus rises; (c) males gain from across-district marriage in districts where marriage migration is high while females gain from within-district marriage in districts where marriage migration is low. This appears to point towards that marriage migration helps men and hurts women.

However, estimating the static marriage market model with transferable utility has several limitations. First, it identifies the marriage surplus based on observed marriage choices, so at the beginning of the marriage. However, the marital gains may vary over the life cycle (Calvi, 2020). Also, the identification of marital gains depends on the number of single observations, which is very low in the case of India; marriage is nearly universal. To overcome those limitations, we propose a collective household model based on Lise and Seitz (2011). This modeling approach enables the identification of partners' bargaining power and overcomes the limitations faced by the static model with transferable utility.

Literature review

This paper is related to several strands of literature. First, we contribute to the literature on marriage migration in India. By analyzing the causes and consequences of moving to their husband's household in another district, we extend the literature that focuses on long-distance marriage migration in India. Kaur (2012, 2013) shows the importance of cultural differences across India for the position of women within the household. In particular, women migrating from other regions face the burden of adjusting to another culture, which results in discrimination and domestic violence. The discrimination relates to i.e. skin color or cultural elements (Chaudhry, 2019). Ahlawat (2009) shows that the consequences go beyond cultural adjustment cost. Moving far away from the parental home negatively affects women's mental health, and the effect is persistent even after several years of marriage. Also, a long distance from the parental home makes them more vulnerable (Kukreja and Kumar, 2013). Further, the discrimination might be fostered by a negative image of long-distance marriage migration generated by the media (Mishra, 2021). Finally, Chaudhry and Mohan (2011) provides evidence that men who search for partners outside the district are negatively selected on landownership, age, prior marital status, or reputation. Using the static marriage market model, we show that men gain in utility terms from across-district marriages in districts where marriage migration is common. In contrast, women derive hither utility from within-district marriages in districts where marriage migration is rare.

By analyzing the spatial correlation between the sex ratio and the intensity of marriage migration, we contribute to the literature that studies the consequences of a skewed sex ratio for the marriage market and women empowerment in India. Anukriti (2013) finds that gender imbalance that results from strong son preferences leads to lower educational attainment, age at marriage, and labor force participation of Indian women. Further, it also increases the age gap between partners. Those correlations suggest a skewed sex ratio might negatively affect women's bargaining power. Foster and Rosenzweig (2001) provides contrary evidence regarding women's employment and suggests that in the regions with more skewed sex ratios, the women's position within the household is higher. Kaur (2008) and Borker et al. (2022) show that the relationship between the sex ratio and the marriage market is not unilateral. The institution of marriage in India includes dowries and the wife's migration to the husband's household. It means the main cost of marriage is attributed to the bride's family. As a result, it generates stronger preferences for sons and daughters-in-law, resulting in a skewed sex ratio (Jayachandran, 2015, Alfano, 2017, Bhalotra et al., 2020). Bhaskar (2011) embodies this idea into the theoretical models and uses it to analyze the role of sex ratio in the marriage market and the role of the marriage market for the abortion of girls and skewed sex ratio. This paper complements the literature by providing empirical evidence linking the sex ratio skewness and marriage migration. Further, we use the model to show that marriage migration impacts marital surplus and how partners split it, which indicates their bargaining power.

Finally, we create a theoretical collective household model which can be used to study the relationship between marriage migration, dowries, and female bargaining power. By that, we contribute to studies focusing on dowries in India. In recent years, the size of dowry in India has increased rapidly (Edlund, 2006). Anderson (2003, 2007), Sautmann (2011) and Rao (1993) attribute the inflation of dowries to the population growth and caste system, while Chiplunkar and Weaver (2020) suggests that the increase in dowries is due to the increase in the quality of grooms. The size of the payment to the groom's family is crucial since it directly impacts female bargaining power. Calvi and Keskar (2021) estimate the collective household model and show that share of consumption allocated to women is strongly associated with the dowry size. Salem (2018) and Brown (2009) come to similar conclusions also for other countries where dowry customs are widely spread. Additionally, Chaudhry and Mohan (2011) and Kaur (2012) show that long-distance marriage migration is often associated with lower or lack of dowry. Our theoretical collective household model provides a framework to test if parents use marriage migration to trade lower dowries for the price of the bride's lower bargaining power in the future.

The rest of the paper is organized as follows. Section 3.2 discusses data and presents stylized facts about sex ratio and marriage migration in India. Next, in Section 3.3, we analyze individual and household characteristics of female marriage migrants. Further, Section 3.4 presents a static marriage market model with transferable utility and its estimation outcomes. In Section 3.5, we build a theoretical collective household model to study the relationship between dowries, marriage migration, and women's bargaining power. Finally, Section 3.6 concludes.

3.2. Stylized facts

In this section, we present statistical evidence exploring the link between sex ratios and marriage migration. First, we describe the datasets used in the analysis. Next, we provide some information about marriage patterns in India. Finally, we show empirical evidence linking sex ratio and probability of marriage migration.

3.2.1. Data sources and sample selection

We use two different data sources to conduct the empirical analyses. The first dataset is the Census of India 2001, a national survey conducted every ten years to gather the information about the Indian population. The second dataset is the National Sample Survey of 2006/07 (NSS), a representative nationwide household survey.

We use the Census for the year 2001 to construct the sex ratio for different 5 year age groups. We express the sex ratio as the number of men per 100 women. In India, the sex ratio is skewed towards men. Therefore, in the rest of the paper, the worsening of sex ratio refers to an increase in the sex ratio. Further, we standardize the sex ratio at the state level. We do so to account for the fact that the variation in sex ratio within the state rather than at the national level drives across-district marriage migration. Next, we divide districts by terciles of standardized sex ratio into three categories: Worse, Neutral, and Better.

The NSS provides information on basic household demographic characteristics and labor activities. We use the 2006/07 survey as it includes detailed information about the migration history of all household members, enabling us to identify women who migrated across districts for marriage purposes. This features allows us to construct the probability of being a marriage migrant. The proportion of individuals who migrated for marriage outside district but within state are referred to as the probability of marriage migration in this paper. We also construct a standardized measure of this variable.

For our analysis, we focus on the age group of 20 to 34 years for females in 2006/2007. Therefore, we construct the sex ratio for the age group of 15-19 years as a measure of marriage market tightness. However, there are certain limitations to this measure. First, as we analyze 20 to 34 years, the measure of tightness is reflective for only part of the cohort. However, Guilmoto and Attané (2007) find no significant improvement in the child sex ratio between 1991 and 2001, suggesting relatively stable over time and across district incentives for having a son. Therefore, the sex ratio for 15-19 years serves as a good proxy for the tightness of the marriage market for the entire age group of 20 to 34 years.

Second, around 25 percent of all marriages occur in the age group of 15 to 19 years. This implies that not only does this measure include some married women but it would also include some women who have migrated for marriage, thus, adding some potential bias. However, as the proportion of marriages is still relatively low, the size of the bias is likely to be small.

We will limit our analysis the the states that have a population of at least 20 million in the year 2001 [2,3]. Table 3.1 presents descriptive statistics for the NSS 2006/07 sample. The top and middle segment of the table provides individual and household characteristics. In terms of individual characteristics, the mean age of the sample is 27 years. Slightly more than half of the women in the sample have primary or higher education. Turning to household characteristics, the average number of children per household is slightly above 2. Less than 1 out of 5 households belongs to the non-Hindu religion group. Most of the households are in rural areas, and more than half of them possess the land. The mean consumption expenditures per capita are around thousand rupees.

3.2.2. Marriage in India

Marriage in India is nearly universal, with only 2.2 percent of women remaining single in the age group of 30 to 34 years (Srinivasan and James, 2015). According to the Census 2001, the average age at marriage for women is 18.3 years and for men is 22.6 years;

[2] Link

[3] The states and UTs that are not included are: Jammu & Kashmir, Himachal Pradesh, Chandigarh, Uttaranchal, Delhi, Sikkim, Arunachal Pradesh, Nagaland, Manipur, Mizoram, Tripura, Meghalaya, Daman & Diu, Dadra & Nagar Haveli, Goa, Lakshadweep, Pondicherry, Andaman & Nicobar Islands.

Table 3.1: Demographic characteristics of the NSS 2006/07 sample

	Mean	Std. dev
Individual:		
Age	26.972	4.052
Primary education or higher	0.511	0.500
Household:		
Number of children	2.188	1.511
Non-Hindu	0.168	0.374
Urban	0.243	0.429
Landowners	0.558	0.497
Consumption per capita (in 1k INR)	0.938	0.681
Marriage migration:		
Migrant	0.731	0.443
Migrant from same state but different district	0.202	0.402
Migrant from another state	0.041	0.198
Observations	49,038	

Notes: 1. The sample consists of married women aged 20-34. Please see Section 3.2.1 for further details. 2. A female is a marriage migrant if they change their place of enumeration due to marriage. 3. All statistics are computed using population weights. *Source:* Authors' calculations using NSS 2006/7.

however, the legal age at marriage for a woman is 18 years and for a man is 21 years [4].

In India, when women get married, they move to their husband's house, which is referred to as the patrilocality of marriage. There are many rules that govern marriage in India: parents often decide and arrange the marriage, ensuring that their children marry within the caste (caste endogamy), language, culture, region and religion. Between North and South India, a key difference is in the practice of village exogamy - in the north of India, females are married into households which are not in the same village (often even in the 2-3 villages in the nearby vicinity), whereas in the south, there are less restrictions on this (Jejeebhoy and Halli, 2005).

The bottom part of Table 3.1 shows statistics regarding marriage migration. In line with previous findings, most Indian women migrated for marriage. Further, 20% of all women migrate to the husband's household within the same state but to another district. It indicates that a relatively high share of Indian women had to leave their district and move to another one to live with their husbands. However, only 4% of women changed to another state due to marriage. It suggests that while crossing the district border within the state for marriage purposes is relatively common practice, moving to another state

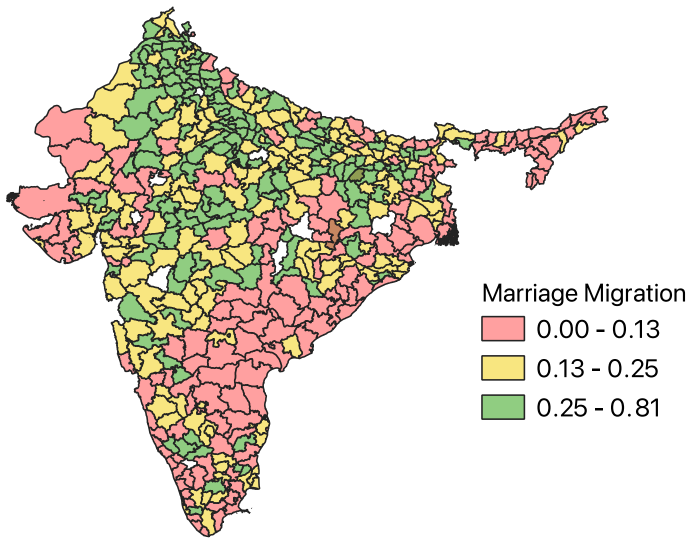
[4]This is a current topic of debate, where the current government plans to raise the age at marriage of females to 21 years, same as males.

for the same reason is still relatively rare.

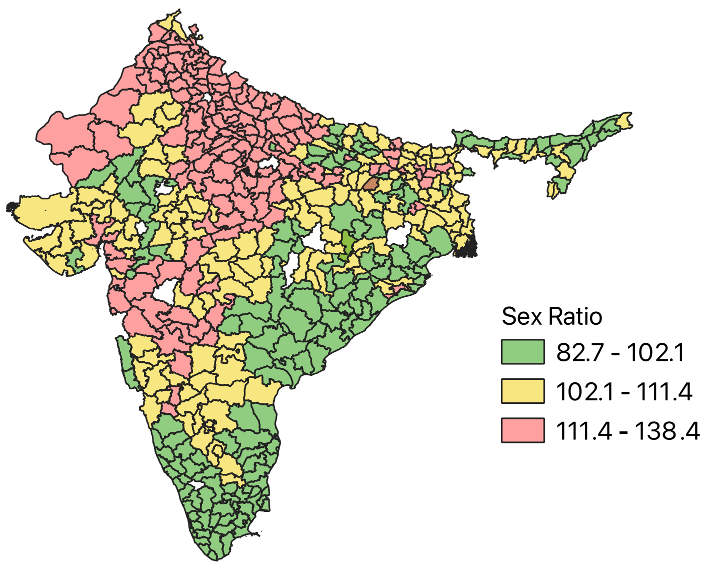
3.2.3. Across-District Marriage Migration in India

Figure 3.1: Sex Ratio and Across-District Marriage Migration in India

(a) Marriage Migration



(b) Sex Ratio



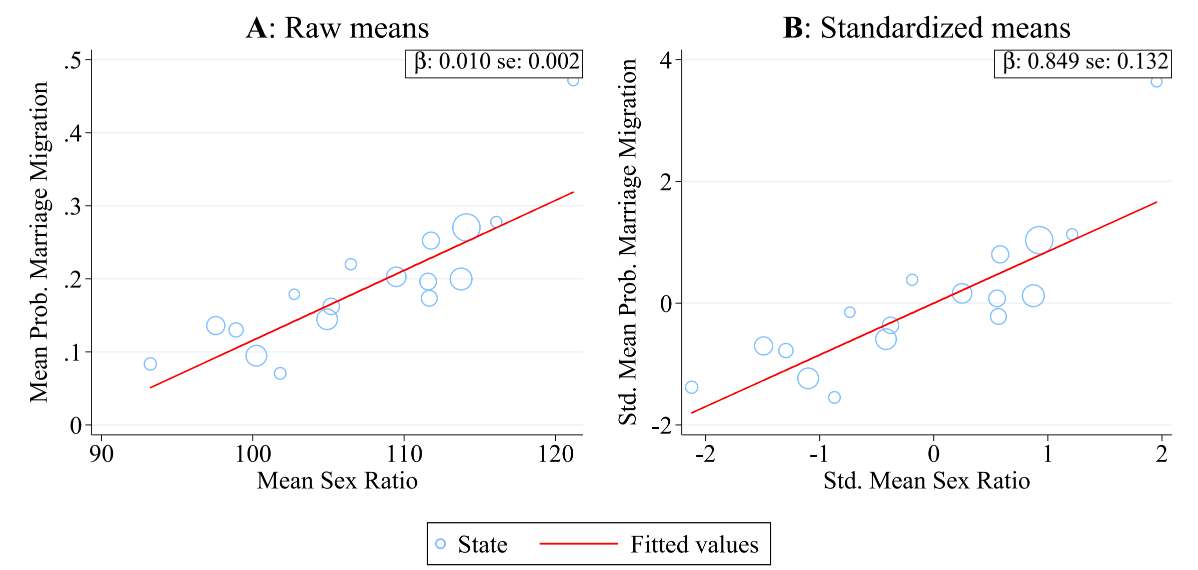
Source: Authors' calculations using Census of India 2001 and NSS 2006/7

Notes: Sex ratio is measured as number of men per 100 women. The proportion of individuals who migrated for marriage outside of their district but within state are referred to as across-district marriage migration.

Marriage migration across districts in India is relatively common, but its intensity varies significantly across different regions. Panel (a) of Figure 3.1 shows the spatial distribution of across-district marriage migration in India. The districts are categorized into three groups based on the intensity of marriage migration: high, moderate, and low. Notably, four districts stand out with over 70 percent of brides migrating across districts: Kannauj (Uttar Pradesh), Bhiwani (Haryana), Kurukshetra (Haryana), and Lohardaga (Jharkhand). These districts are located in the northern states, where a high level of marriage migration is observed in most districts. Conversely, across-district marriage migration is relatively less common in the southern and western parts of India, with some districts reporting nearly no instances or very rare occurrences of such migration.

Across-district marriage migration can be influenced by various factors, and one possible factor is the variation in the sex ratio. Districts with a significantly higher number of men compared to women in the marriageable population may be a migration destination for women. In such districts, the marriage market tends to be more competitive, leading men to seek partners outside their own district. Panel (b) of Figure 3.1 illustrates the distribution of the sex ratio across different regions. Regions with highly skewed sex ratios are concentrated in the northern part of the country, while relatively balanced sex ratios are observed in the southwestern part of India. There seems to be a correlation between the sex ratio distribution and the proportion of across-district marriage migrants. Specifically, districts with a high sex ratio tend to have a higher prevalence of marriage migration as well.

Figure 3.2: Variation of Migration with Sex Ratio, State Level



Source: Authors' calculations from Census 2001 and NSS 2006/7. *Notes:* 1. Sex ratio is measured as number of males per 100 females. 2. Standardized values are constructed within states.

However, this correlation might be spurious. It might be a case that districts within states with very skewed sex ratio are characterised by the high marriage migration, but not necessarily women move from low to high sex ratio districts. To verify this hypothesis, we focus on the relative sex ratio within a states. Therefore, focusing on districts with Better and Neutral sex ratio (within state), the across-district marriage migration is 19 percent, on average; this number rises to 26 percent for those with Worse sex ratio. Focusing on the correlation between standardized marriage migration and standardized sex ratio, the proportion of marriage migration within district increases by 0.13 standard deviations, for each standard deviation increase in the sex ratio. In other words, as the sex ratio worsens, the proportion of marriage migration increases. This implies that there is significant correlation of the sex ratio with marriage migration, *within state*.

Figure 3.2 presents the correlation of the levels of sex ratio with the probability of marriage migration, both levels (panel A) and standardized (panel B), at the state level. We find a positive relationship: as the sex ratio worsens, the average probability of marriage migration increases, *across states*. Further, for each standard deviation increase in the sex ratio, there is a 0.849 standard deviation increase in the probability of marriage migration. This is captured in the previous graph as well as we see a bunching of states by their type with Worse sex ratio states on the higher end of the distribution.

Thus, from the two graphs above, we show that across district migration increases with sex ratio, within and across states.

3.3. Individual and household characteristics of marriage migrants

We now analyse the determinants of marriage migration in India using a reduced form approach. Due to the data availability, we focus on individual characteristics of women and characteristics of the household they marry into. This allows us to understand what distinguishes women who leave their district because of marriage and which type of households are more likely to attract them. We focus on married women who choose to migrate within state[5]. Therefore, the dependent variable y in our analysis is 1 for those who migrate outside the district but within the state, 0 otherwise.

Formally, the probability that woman f who lives in district d is a marriage migrant from another district but within the same state is:

$$\mathrm{P}(y_{id} = 1) = L\left(\alpha_0 + \alpha_1 \log(sr_{id}) + X_i\beta + Z_i\gamma + \delta_{s(d)}\right) \tag{3.1}$$

Here, $L()$ represents the logistic transformation. The coefficient α_1 quantifies the impact of the sex ratio in district d, denoted by sr_{id}, on the likelihood of marriage migration. The vector X_i contains individual characteristics such as age and education, while Z_i contains household characteristics including the educational level of the household head, location of household (urban/rural), social group, household consumption, and landownership. Additionally, we control for state fixed effects denoted by $\delta_{s(d)}$.

Table 3.2 presents estimated coefficients for two models given by Equation 3.1 with

[5]In this analysis, we drop married women who migrated outside the state - these is 4.1 percent of married women in the age group of 20-35 years.

Table 3.2: Logistic Regression for Marriage Migration

	All states	All states
Log of sex ratio	3.622***	1.055***
	(0.182)	(0.227)
Individual characteristics:		
Age	-0.001	-0.001
	(0.004)	(0.004)
Primary educ. or higher	0.226***	0.261***
	(0.038)	(0.039)
Household characteristics:		
Head with primary educ. or higher	0.163***	0.147***
	(0.037)	(0.037)
Urban	-0.084*	-0.030
	(0.044)	(0.045)
Scheduled Tribe	-0.294***	-0.346***
	(0.069)	(0.073)
Scheduled Caste	0.065	-0.009
	(0.048)	(0.050)
Other backward classes	0.106***	0.056
	(0.039)	(0.043)
Log of consumption per capita	0.054	0.096**
	(0.039)	(0.041)
Landowners	0.013	-0.049
	(0.035)	(0.037)
Observations	51,775	51,775
State FE	No	Yes

Notes: 1. The sample consists of all women aged 20-34 who did not move out of their birth state. 2. Sex ratio is the number of men per 100 women in the youngest tercile of marriageable group in the current district. 3. Household consumption is defined as monthly household expenditures per capita in thousands of Rupees. 4. Household owns land if the acreage possessed is more than 0.01 hectares. *Source:* Authors' calculations from the NSS 2006/7. $*p < 0.1, **p < 0.05, ***p < 0.01$.

and without state-fixed effects. First, across-district marriage migrants are more likely to have at least primary education, and they move to households where the head is also more likely to have at least primary education. Second, rural households are more likely to search for brides outside their districts. This coefficient, however, is significant only when we do not control for state-fixed effects, indicating that states with high urbanization drive the initial effect. Third, scheduled tribe households are less likely to have across-district spouses. Fourth, wealthier households or households with higher consumption expenditures per capita are more likely to have spouses from outside the district. The coefficient is positive for both specifications, but the significance holds only while controlling for state-fixed effects. Lastly, as expected, the log of sex ratio in the district the married woman migrates to is positively correlated with the married woman choosing to migrate outside the district. This suggests that marriage migration occurs when the sex ratio is skewed against females. Controlling for state-fixed effects certainly dampens the effect, which implies that marriage migration is more common in the states which are characterized by high levels of sex ratio, as has been seen in our previous results.

We also conduct additional analyzes, where we split the sample into two groups: women living in the states with a less skewed sex ratio (Better) and a more skewed sex ratio (Worse). Table C.1 presents these results. The sign of most of the coefficients is similar in both cases. However, there are significant differences in the magnitudes. First, for worse sex ratio states, if both woman and head have primary education or higher, this results in higher marriage migration probability as compared to better sex ratio states where the significant effect comes from only the woman's education. Second, in the better sex-ratio states, household belonging to scheduled tribes are less likely to search for a woman outside of their district than in the worse sex-ratio states. The opposite conclusion holds for the backward classes. Finally, the sex ratio seems to play a stronger role in the states with a more skewed sex ratio. It suggests that in the worse sex ratio states, the correlation between the spatial distribution of sex ratio and marriage migration is even stronger than in the other states.

Finally, we further analyze our sample by dividing it into women residing in urban and rural areas. Table C.2 presents the estimated coefficients for these two groups. In urban areas, the education of women and the household head appear to have a positive influence on the probability of marriage migration. However, when it comes to wealth measured by consumption expenditure, it does not significantly affect the probability in urban areas. In contrast, it has a statistically significant impact in rural areas. This suggests wealthier households in rural areas are more likely to search for a bride in another district. Moreover, the district sex ratio plays a stronger role in rural areas. This supports the intuition that households choose to search for partners in other districts when the local marriage market is relatively smaller and, as a result, households face greater challenges due to its tightness.

3.4. Model with transferable utility

This section presents the first modeling approach to investigate consequences of marriage migration in India. We develop the simple mode that is based on seminal work by Choo and Siow (2006). Next, we present outcomes of the estimation and briefly discuss the results and its limitations.

3.4.1. Model description

In this model, we consider a society consisting of males (m) and two types of females: those searching for a partner within-district (f_1) and those searching for a partner across-district (f_2). The choice of types is driven by the data suggesting that while a significant proportion of females migrate for marriage, very few males do. As a result, the model only allows for female migration.

Since our primary focus is on within-district migration, we detail the model for a single state. Each individual maximizes their utility by choosing whom to marry or whether to remain single. The number of men married to type f_i women is denoted as μ_{mf_i}. We denote the number of unmarried men (women) of type m (f_i) as μ_{m0} (μ_{0f}). The market clearing conditions are given by:

$$\mu_{m0} + \mu_{mf_1} + \mu_{mf_2} = |m| \tag{3.2}$$

$$\mu_{0f_1} + \mu_{mf_1} = |f_1| \tag{3.3}$$

$$\mu_{0f_2} + \mu_{mf_2} = |f_2| \tag{3.4}$$

$$\mu_{m0}, \mu_{0f}, \mu_{mf_1}, \mu_{mf_2} \geq 0 \tag{3.5}$$

The model is built on the assumption of transferable utility (TU), which means that agents can transfer part of their utility to their partner in equilibrium. The individual utility of a man g who marries a type i woman is given by:

$$V_{gf_i} = \tilde{\alpha}_{mf_i} - \tau_{mf_i} + \varepsilon_{gf_i} \tag{3.6}$$

Similarly, the utility of a woman h of type i who marries a man is given by:

$$U_{hf_im} = \tilde{\gamma}_{mf_i} + \tau_{mf_i} + \eta_{hf_im}, \tag{3.7}$$

Here, α_{mf_i} and γ_{mf_i} represent the gross return for a man married to a type i woman and a type f_i woman married to a man, respectively. τ_{mf_i} is the equilibrium transfer from a man to his type f_i spouse, and ε_{gf_i} and η_{hf_im} are independently and identically distributed shocks following a type I extreme value distribution.

Agents can also choose to remain single. The utility of a single woman h and a single man g (type $f_i = 0$) is given by:

$$V_{g0} = \tilde{\alpha}_{m0} + \varepsilon_{g0} \tag{3.8}$$
$$U_{hf_i0} = \tilde{\gamma}_{0f_i} + \eta_{hf_i0}, \tag{3.9}$$

Here, ε_{g0} and η_{hf_i0} are also independently and identically distributed shocks following a type I extreme value distribution.

Using the assumptions on the distribution of utility shocks and the fact that agents choose their partners to maximize their utility, we can derive the quasi demand equations for men and women:

$$\ln \mu^D_{f_im} = \ln \mu_{m0} + \alpha_{mf_i} - \tau_{mf_i} \tag{3.10}$$
$$\ln \mu^S_{f_im} = \ln \mu_{0f_i} + \gamma_{mf_i} + \tau_{mf_i}, \tag{3.11}$$

Given equilibrium transfers τ_{mf_i}, the demand by men for a type f_i women is equal to supply of a type f_i women for a men for all possible combinations of women types. α_{mf_i} (γ_{mf_i}) represents the systematic gross return for a man (type f_i woman) from an (m, f_i) marriage relative to being unmarried.

Combining the quasi demand and quasi supply equations, we have:

$$\ln \mu_{mf_1} - \frac{\ln \mu_{m0} + \ln \mu_{0f_1}}{2} = \frac{\alpha_{mf_1} + \gamma_{mf_1}}{2} \tag{3.12}$$
$$\ln \mu_{mf_2} - \frac{\ln \mu_{m0} + \ln \mu_{0f_2}}{2} = \frac{\alpha_{mf_2} + \gamma_{mf_2}}{2} \tag{3.13}$$

We refer to $\frac{\alpha_{mf_1}+\gamma_{mf_1}}{2}$ and $\frac{\alpha_{mf_2}+\gamma_{mf_2}}{2}$ as the total marriage surplus within- and across-district, respectively.

In addition to the marriage surplus, Equations (3.10) and (3.11) can be used to identify $\alpha_{mf_i} - \tau_{mf_i}$ and $\gamma_{mf_i} + \tau_{mf_i}$:

$$\ln \frac{\mu_{mf_1}}{\mu_{m0}} = \alpha_{mf_1} - \tau_{mf_1} \tag{3.14}$$
$$\ln \frac{\mu_{mf_2}}{\mu_{m0}} = \alpha_{mf_2} - \tau_{mf_2} \tag{3.15}$$
$$\ln \frac{\mu_{mf_1}}{\mu_{0f_1}} = \gamma_{mf_1} + \tau_{mf_1} \tag{3.16}$$
$$\ln \frac{\mu_{mf_2}}{\mu_{0f_2}} = \gamma_{mf_2} + \tau_{mf_2} \tag{3.17}$$

Equation (3.14) and (3.15) represent the systematic net gain to marriage for a man in a marriage with f_i relative to remaining single. Similarly, Equation (3.16) and (3.17) define the systematic net gain to marriage for a woman of type f_i relative to remaining single.

3.4.2. Model results

The estimation of the marriage surplus and net gains, as defined by Equations (3.10)-(3.17), requires data on the number of married and single individuals of each type. While obtaining the numbers for married individuals is straightforward, determining the number of single females searching for a husband across districts is challenging due to data limitations. To address this issue, we construct the number of across-district singles as a weighted average of all females in districts within the state, excluding the district under consideration. For instance, in a state with three districts, for district 1, μ_{0f_2} would be the weighted average of single females in districts 2 and 3.

Figure 3.3 illustrates the estimated total marriage surplus for within-district marriages (a) and across-district marriages (b). Firstly, within districts, the marriage surplus is lower in the northern compared to the western and south-eastern regions. However, this pattern is reversed for across-district marriage surpluses, where the North exhibits significantly higher surpluses. Secondly, this distribution of marriage surplus is correlated with the sex ratio and marriage migration. Specifically, areas with a high sex ratio and high marriage migration tend to have lower within-district surpluses but higher across-district surpluses.

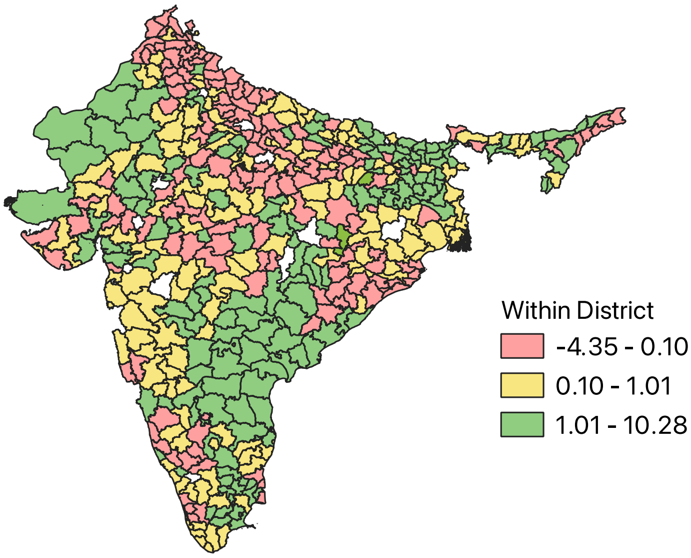
In Figure 3.4, we examine the relationship between marriage surplus and marriage migration. The probability is standardized at the state level, and the size of each circle represents the population size of the corresponding district. The districts are categorized into three groups based on their sex ratio, as explained in Section 3.2.

The results reveal a negative and statistically significant correlation between within-district marriage surplus and marriage migration. Nevertheless, a positive relationship is observed between across-district marriage surplus and marriage migration, although this relationship is not statistically significant. These findings suggest that in regions where the overall surplus from marrying within the same district is low or where the surplus from across-district marriages is high, men are more likely to seek partners outside their own district.

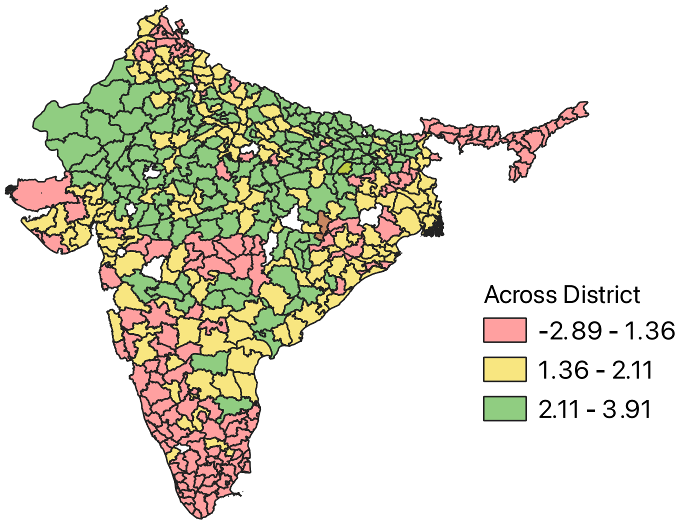
Figure 3.5 illustrates the correlation between division of the total marriage surplus between men and women, as determined by Equations 3.14 to 3.17, and probability of marriage migration. The results suggest a negative correlation between gains from marriage for men and the probability of marrying a woman within the same district. Conversely, there is a positive correlation, albeit statistically insignificant, between gains from marriage for men and the probability of marrying a woman from outside their district. Furthermore, both types of marriages exhibit a negative correlation with marriage migration in terms of gains among women. This indicates that men benefit from marrying across districts in regions with high levels of marriage migration. Nevertheless, women benefit from marrying within their own district in areas with low levels of marriage migration. This finding aligns with the notion that districts with low marriage migration tend to have a more favorable sex ratio.

Figure 3.3: Total Marriage Surplus Across and Within District in India

(a) Within-District



(b) Across-District

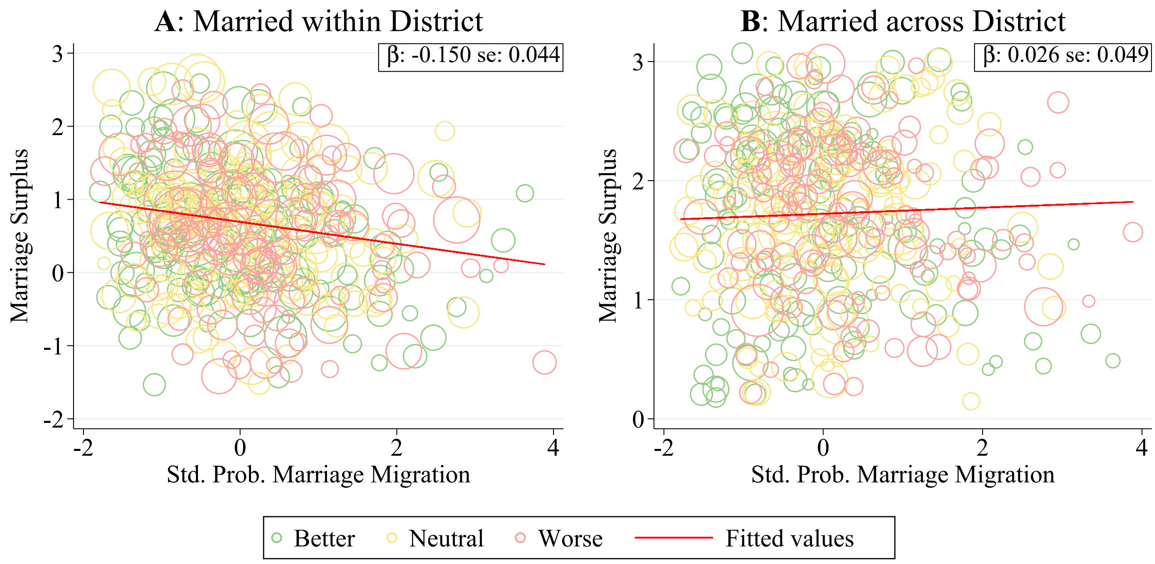


Source: Authors' estimation using NSS 2006/7. Please see Section 3.4.1 for further details on estimation.

Analyzing the relationship between marital surplus and marriage migration at the state level (Figure 3.6) strengthens the effect seen at a district level that states with a low mean probability of marriage migration are characterized by the high total within-district marriage surplus and the low across-district marriage surplus.

In summary, there are three key takeaways from the model: (1) northern states have a higher across-district surplus, while south-eastern and western states have a higher within-district surplus. (2) As marriage migration increases, within-district marriage

Figure 3.4: Variation of Marital Surplus with Marriage Migration



Source: Authors' calculations from Census 2001 and NSS 2006/7. *Notes:* 1. Sex ratio is measured as number of men per 100 women. 2. Standardized values are constructed within states. 3. Marriage surplus trimmed on the graph to 5th and 95th percentile. 4. Size of the circle represents district's population.

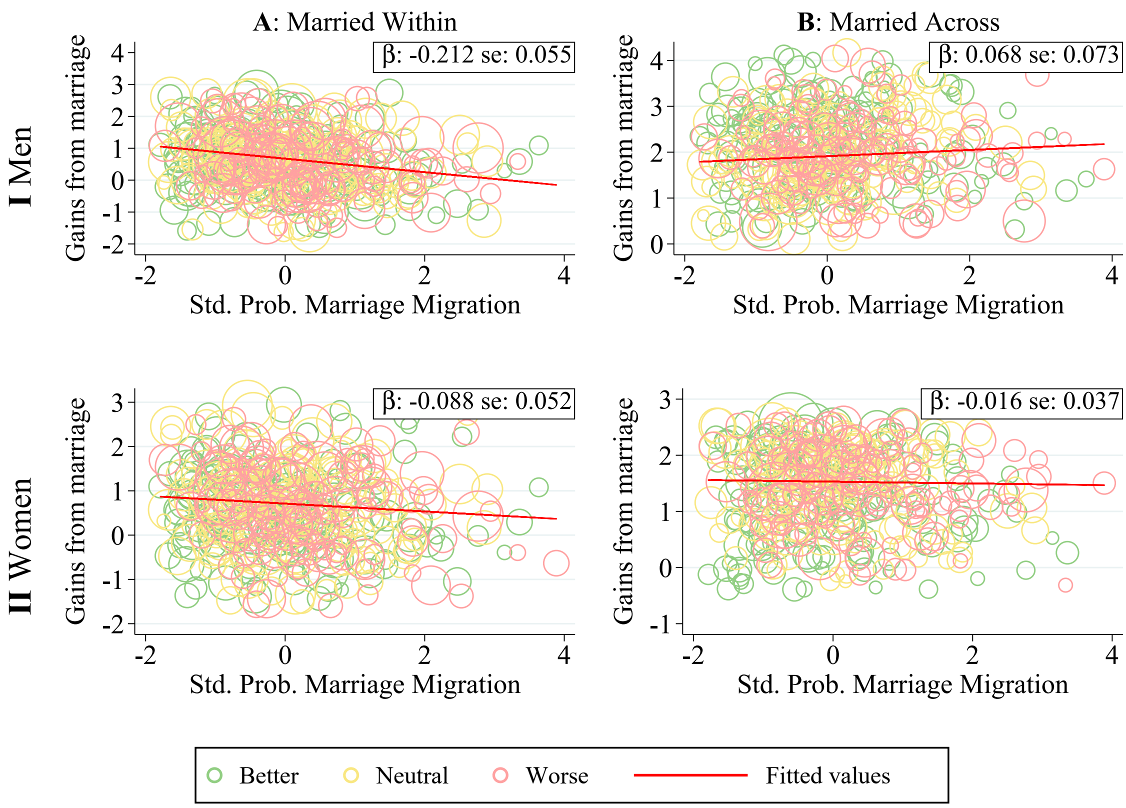
surplus falls, and across-district surplus increases, both at the district and state levels. (3) Males gain from across-district marriage in districts where marriage migration is high, while females gain from within-district marriage in districts where marriage migration is low. This appears to point towards that marriage migration helps men and hurts women.

3.4.3. Limitations

The proposed modelling approach faces several limitations. First, there is a low number of single individuals in the dataset. As mentioned in Section 3.2.2, marriage is almost universal in India and this can have implications for cleanly identifying the surplus. Second, we do not distinguish between types of men or women, in terms of education and age. This means that we assume that marital gains (potential measures of the bargaining power) remain the same for all women and men, and varies only by migration status. Consequently, the outcomes of the model are homogeneous, and it becomes difficult to identify which women are genuinely affected by the challenges of marriage migration. Lastly, the identification of the bargaining power is based on marriage choices. However, it might be the case that bargaining power changes with marriage duration (Calvi, 2020). As a result, the estimated marital gains do not accurately summarize the life-time gains.

To address these limitations and gain more detailed insights, a different modeling approach is needed. This alternative model should allow for the identification of bargaining power at the household level and establish stronger links between bargaining power

Figure 3.5: Variation of Individual Gains from Marital Surplus with Marriage Migration



Source: Authors' calculations from Census 2001 and NSS 2006/7. *Notes:* 1. Sex ratio is measured as number of males per 100 females. 2. Standardized values are constructed within states. 3. Marital gains trimmed on the graph to 5th and 95th percentile. 4. Size of the circle represents district's population.

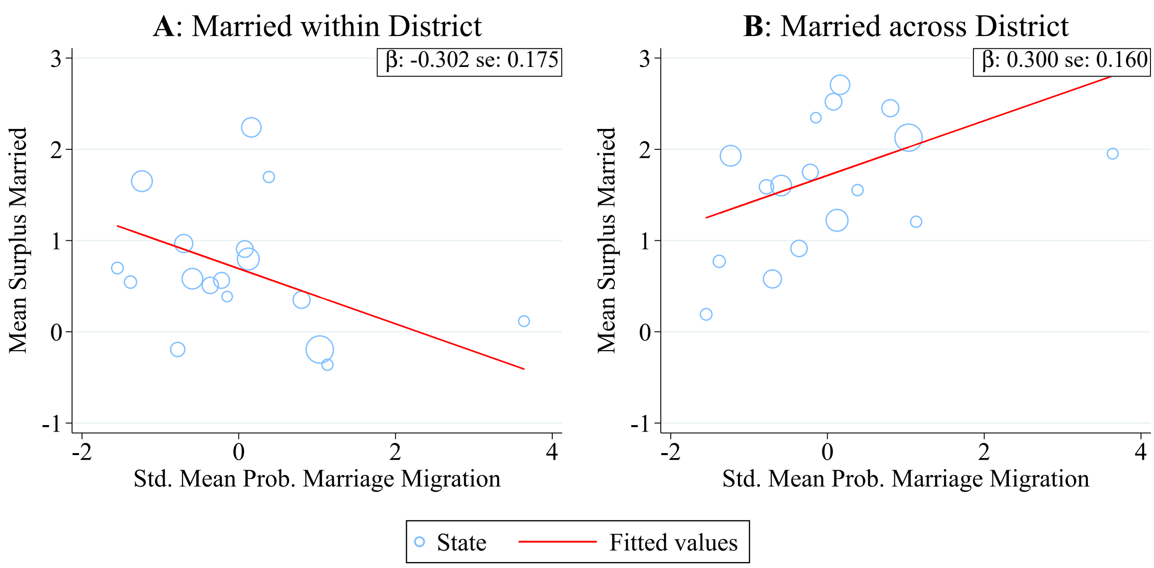
and marriage migration. Such a model would offer a more comprehensive understanding of the dynamics involved, shedding light on the nuanced experiences of individuals and providing a more accurate assessment of the impacts of marriage migration. We present a version in the next section.

3.5. Collective household model

In this section, we briefly describe the collective household model that can be used to identify sharing rules within households. These sharing rules are related to partners' bargaining power and are commonly employed as a measure of intra-household gender inequality Browning et al. (2013).

The theoretical model we utilize to identify bargaining power within the household is based on the collective model of household decision-making proposed by Lise and Seitz (2011). This model is able to address several key issues. Firstly, it enables the identification of Pareto weights and the creation of a measure of bargaining power. Secondly, it

Figure 3.6: Variation of Marital Surplus with Marriage Migration, State Level



Source: Authors' calculations from Census 2001 and NSS 2006/7. *Notes:* 1. Sex ratio is measured as number of males per 100 females. 2. Standardized values are constructed within states. 3. Size of the circle represents states's population.

distinguishes between private and public consumption, which is crucial as ignoring private consumption may lead to biased measures of bargaining power within households. Lastly, the model allows for the identification of the location of the sharing rule, extending the standard identification beyond an additive constant Chiappori (1988).

3.5.1. Theoretical framework

We consider households with only tow decision makers:: man m and women d. Each partner $j \in \{m, f\}$ has a distinct preference over own leisure ℓ_j, own private consumption c_j and household public consumption C. We assume that preferences over private consumption and leisure are separable from consumption of the public good. The household budget consists of non-labor income y_{nl} and labor income that depends on partners' labor supply choices. Partner j chooses optimal working hours h_j given hourly wage w_j.

Under assumptions that preferences are egoistic and that allocations are Pareto efficient, the household maximization problem is as follows:

$$\max_{c_m, c_d, C, \ell_m, \ell_m} \mu(\pi, y, \boldsymbol{z}) U_m(u_m(c_m, \ell_m), C) + (1 - \mu(\pi, y, \boldsymbol{z})) U_f(u_f(c_f, \ell_f), C) \tag{3.18}$$

under the budget constraint:

$$c_m + c_f + pC + \ell_m w_m + \ell_f w_f = T w_m + T w_f + y_{nl} = y \tag{3.19}$$

where T is time endowment and $\mu(\pi, y, \boldsymbol{z})$ is the sharing rule that depends on prices π, total resources y and distribution factors $\boldsymbol{z}$. The budget constraint is defined in the way that the expenditures on consumption and leisure are on the left hand-side and full potential income is on the right hand-side.

Following Blundell et al. (2005), we decentralized household allocation problem into a two-stage process. This is possible as we assume that leisure and private good consumption are separable from the public good consumption.

In the first stage, partners take two decisions. They decide on household public consumption C. They also agree on a particular distribution of the non-labor income net of public consumption expenditures between them. This takes a form of a transfer from man m to his wife f. The transfer can be negative or positive. In the second stage, each household member freely chooses their level of labor supply and private consumption subject to the budget constraint stemming from the first stage.

Formally, the woman's problem in the second stage is as follows:

$$\max_{c_f, \ell_f} u_f(c_f, \ell_f) \tag{3.20}$$

such that:

$$c_F + \ell_f w_f = T w_f + \psi(w_m, w_f, y_{nl}, \boldsymbol{z}) \tag{3.21}$$

The man m's problem in the second stage is as follows:

$$\max_{c_m, \ell_m} u_m(c_m, \ell_m) \tag{3.22}$$

such that:

$$c_m + \ell_m w_m = T w_m - \psi(w_M, w_F, y, \boldsymbol{z}) \tag{3.23}$$

Here, $\psi(w_m, w_f, y_{nl}, \boldsymbol{z})$ define the conditional transfer between partners. The transfer is conditional since partners share non-labor income net of public good expenditures. It is important to note here that the size of the transfer is a function of partner potential earnings, household non-labor income and distribution factors. It means that it does not depend on the labor supply choice of either partner.

3.5.2. Sharing rule and marriage migration

The transfers between partners serve as a measure for the bargaining power within the household. We use this measure to study the effect of exposure to marriage migration on women's position in the household.

First, we define $\rho_{fm} \in [0, 1]$ as a share of household full income transferred from man m to woman f. Next, we denote by P_f the probability that woman f is a marriage migrant. The probability of marriage migration is calculated outside of the collective model, and is defined by the Equation 3.1. Existing literature finds that woman who migrated long-distance on average might suffer from the distance to parental home and do not have support of their family (Kaur, 2012, Kukreja and Kumar, 2013, Mishra, 2021). Nevertheless, they might gain some bargaining power due to the fact that they move to the region with more skewed sex ratio, where women are in scare number and thus, might be more favorable for them. With our model we can test this hypothesis. Formally, the the null hypothesis is:

H0: $\text{cov}(\rho_{mf}, P_f) < 0$ - There is a negative relationship between bargaining power of women and the probability of being a marriage migrant.

Second, the literature also suggest that there is a correlation between women's bargaining power and size of a dowry (Calvi, 2020, Calvi and Keskar, 2021, Salem, 2018, Brown, 2009, Anderson and Bidner, 2015). We denote by D_{mf} as a size of a dowry paid by woman f's family to the man m family. Then, the null hypothesis is:

H0: $\text{cov}(\rho_{mf}, D_{mf}) < 0$ - There is a negative relationship between woman's bargaining power and the size of a dowry paid by her family.

Finally, the literature suggests that marriage migration is a common way to avoid dowry or decrease its size by parents of the bride (Chaudhry and Mohan, 2011, Kukreja and Kumar, 2013). Using the data on the size of the dowry and probability of marriage migration, we can also test this hypothesis. Formally:

H0: $\text{cov}(D_{mf}, P_f) < 0$ - There is a negative relationship between woman's dowry and the probability of being marriage migrant.

Testing for these hypotheses allows us to identify the possible trade-off between dowry and women bargaining power. More importantly, if dowry can be decreased by sending the daughter further away from the household, then marriage migration might be seen as a mechanism that allows poorer household to avoid significant expenses associated with the marriage of their daughter. However, it might come with a price of lower position of the daughter in the new household.

3.6. Conclusion

In this paper, we investigate the consequences of across-district marriage migration in India for within-household inequality and women's bargaining power. By analyzing the relationship between migration patterns and the sex ratio, we find a positive correlation between marriage migration and the sex ratio at the district level. A reduced form approach through logistic regression reveals that women with at least primary education are more likely to be marriage migrants. We also find that receiving households are more likely to have a head with at least primary education and, on average, have higher per capita consumption expenditures.

Using a static marriage market model with transferable utility, we examine the effects of marriage migration on the division of marriage surplus between men and women. Our findings indicate that within-district marriage surplus declines with an increase in marriage migration, while across-district marriage surplus rises both at the district and state levels. Further, we find that men benefit from across-district marriage in regions with high migration, while women benefit from within-district marriage in regions where migration is low. However, the static model has limitations, such as the reliance on observed marriage choices and the assumption of constant marital gains over the life cycle.

To address these limitations, we propose a theoretical collective household model that allows identifying bargaining power at the household level. This model is a tool to test hypotheses regarding the correlation between bargaining power and the probability of marriage migration, the relationship between bargaining power and dowry size, and the association between dowry and the probability of marriage migration. By analyzing these relationships, we gain insights into the trade-offs faced by marriage migrants and the impact of marriage migration on women's position within households.

In conclusion, our study contributes to the literature on marriage migration in India by examining its causes and consequences. We provide empirical evidence on the correlation between migration patterns and the sex ratio. Moreover, our models offer insights into the effects of marriage migration on the division of marital gains and women's bargaining power. Future research could further conduct an empirical estimation of the proposed theoretical model. It allows for empirical tests on trade-offs between dowries, marriage migration, and bargaining power.

BIBLIOGRAPHY

Adda, J., Pinotti, P., and Tura, G. (2020). There's More to Marriage than Love: The Effect of Legal Status and Cultural Distance on Intermarriages and Separations. *CEPR Discussion Papers*, (November).

Agostinelli, F., Doepke, M., Sorrenti, G., and Zilibotti, F. (2023). It Takes a Village: The Economics of Parenting with Neighborhood and Peer Effects. *SSRN Electronic Journal*.

Ahlawat, N. (2009). Missing brides in rural Haryana: A study of adverse sex ratio, poverty and addiction. *Social Change*, 39(1):46–63.

Alfano, M. (2017). Daughters, dowries, deliveries: The effect of marital payments on fertility choices in India. *Journal of Development Economics*, 125(November 2016):89–104.

Almond, D., Currie, J., and Duque, V. (2018). Childhood circumstances and adult outcomes: Act II. *Journal of Economic Literature*, 56:1360–1446.

Anderson, S. (2003). Why Dowry Payments Declined with Modernization in Europe but Are Rising in India. *Journal of Political Economy*, 111(2):269–310.

Anderson, S. (2007). Why the marriage squeeze cannot cause dowry inflation. *Journal of Economic Theory*, 137(1):140–152.

Anderson, S. and Bidner, C. (2015). Property rights over marital transfers. *Quarterly Journal of Economics*, 130(3):1421–1484.

Anukriti, S. (2013). Sex Ratios, Schooling, and the Marriage Market in India.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *The Review of Economic Studies*, 58(2):277.

Attanasio, O., Blundell, R., Conti, G., and Mason, G. (2020a). Inequality in socio-emotional skills: A cross-cohort comparison. *Journal of Public Economics*, 191:104171.

Attanasio, O., Cattan, S., Fitzsimons, E., Meghir, C., and Rubio-Codina, M. (2020b). Estimating the production function for human capital: Results from a randomized controlled trial in Colombia. *American Economic Review*, 110(1):48–85.

Attanasio, O., Meghir, C., and Nix, E. (2020c). Human Capital Development and Parental Investment in India. *The Review of Economic Studies*, 87(6):2511–2541.

Aurino, E., Fledderjohann, J., and Vellakkal, S. (2019). Inequalities in adolescent learning: Does the timing and persistence of food insecurity at home matter? *Economics of Education Review*, 70(09):94–108.

Basu, S. (2015). Intermarriage and the labor market outcomes of Asian women. *Economic Inquiry*, 53(4):1718–1734.

Baumrind, D. (1967). Child care practices anteceding three patterns of preschool behavior. *Genetic Psychology Monographs*, 75:43–88.

Becker, G. S. (1973). A Theory of Marriage: Part I. *Journal of Political Economy*, 81(4):813–846.

Becker, G. S. (1974). A Theory of Marriage: Part II. *Journal of Political Economy*, 82(2, Part 2):S11–S26.

Beham, B., Drobnič, S., Präg, P., Baierl, A., and Eckner, J. (2019). Part-time work and gender inequality in Europe: a comparative analysis of satisfaction with work–life balance. *European Societies*, 21(3):378–402.

Benson, S. (1981). *Ambiguous Ethnicity: Inter-racial Families in London.* Cambridge University Press, Cambridge.

Bhalotra, S., Chakravarty, A., and Gulesci, S. (2020). The price of gold: Dowry and death in India. *Journal of Development Economics*, 143(July 2019):102413.

Bhaskar, V. (2011). Sex Selection and Gender Balance. *American Economic Journal: Microeconomics*, 3(1):214–244.

Bloomfield, L. and Kendall, S. (2012). Parenting self-efficacy, parenting stress and child behaviour before and after a parenting programme. *Primary Health Care Research & Development*, 13(04):364–372.

Blundell, R., Chiappori, P., and Meghir, C. (2005). Collective Labor Supply with Children. *Journal of Political Economy*, 113(6):1277–1306.

Boeri, T., De Philippis, M., Patacchini, E., and Pellizzari, M. (2015). Immigration, Housing Discrimination and Employment. *Economic Journal*, 125(586):F82–F114.

Borjas, G. J. (1985). Assimilation, Changes in Cohort Quality, and the Earnings of Immigrants. *Journal of Labor Economics*, 3(4):463–489.

Borker, G., Eeckhout, J., Luke, N., Minz, S., Munshi, K., and Swaminathan, S. (2022). Wealth, Marriage, and Sex Selection.

Brown, P. H. (2009). Dowry and intrahousehold bargaining: Evidence from China. *Journal of Human Resources*, 44(1):25–46.

Browning, M., Chiappori, P. A., and Lewbel, A. (2013). Estimating consumption economies of scale, adult equivalence scales, and household bargaining power. *Review of Economic Studies*, 80(4):1267–1303.

Buettner, T., Erbe, K., and Grimm, V. (2019). Tax planning of married couples and intra-household income inequality. *Journal of Public Economics*, 179(3):104048.

Calvi, R. (2020). Why are older women missing in india? The age profile of bargaining power and poverty. *Journal of Political Economy*, 128(7):2453–2501.

Calvi, R. and Keskar, A. (2021). Dowries, resource allocation, and poverty. *Journal of Economic Behavior and Organization*, 192:268–303.

Calvo, P., Lindenlaub, I., and Reynoso, A. (2021). Marriage Market and Labor Market Sorting. Technical Report July, National Bureau of Economic Research, Cambridge, MA.

Card, D. (1999). Chapter 30 The causal effect of education on earnings. *Handbook of Labor Economics*, 3 PART(1):1801–1863.

Carneiro, P., Crawford, C., and Goodman, A. (2007). The Impact of Early Cognitive and Non-Cognitive Skills on Later Outcomes. Technical report.

Caucutt, E., Lochner, L., Mullins, J., and Park, Y. (2020). Child Skill Production: Accounting for Parental and Market-Based Time and Goods Investments. Technical report, National Bureau of Economic Research, Cambridge, MA.

Chaudhry, S. (2019). 'Flexible' caste boundaries: cross-regional marriage as mixed marriage in rural north India. *Contemporary South Asia*, 27(2):214–228.

Chaudhry, S. and Mohan, T. D. (2011). Of marriage and migration: Bengali and bihari brides in a U.P. village. *Indian Journal of Gender Studies*, 18(3):311–340.

Chiappori, P.-A. (1988). Rational Household Labor Supply. *Econometrica*, 56(1):63–90.

Chiappori, P.-A., Dias, M. C., and Meghir, C. (2018). The Marriage Market, Labor Supply, and Education Choice. *Journal of Political Economy*, 126(S1):S26–S72.

Chiappori, P.-A. and Gugl, E. (2020). Transferable utility and demand functions. *Theoretical Economics*, 15(4):1307–1333.

Chiappori, P.-A. and Mazzocco, M. (2017). Static and Intertemporal Household Decisions. *Journal of Economic Literature*, 55(3):985–1045.

Chiplunkar, G. and Weaver, J. (2020). Marriage Markets and the Rise of Dowry in India. *SSRN Electronic Journal*.

Chiswick, B. R. and Houseworth, C. (2011). Ethnic intermarriage among immigrants: Human capital and assortative mating. *Review of Economics of the Household*, 9(2):149–180.

Chiswick, B. R. and Miller, P. W. (2009). The international transferability of immigrants' human capital. *Economics of Education Review*, 28(2):162–169.

Choi, K. H. and Tienda, M. (2017). Marriage-Market Constraints and Mate-Selection Behavior: Racial, Ethnic, and Gender Differences in Intermarriage. *Journal of Marriage and Family*, 79(2):301–317.

Choo, E. and Siow, A. (2006). Who marries whom and why. *Journal of Political Economy*, 114(1):175–202.

Christensen, D., Fahey, M. T., Giallo, R., and Hancock, K. J. (2017). Longitudinal trajectories of mental health in Australian children aged 4-5 to 14-15 years. *PLOS ONE*, 12(11):e0187974.

Cobb-Clark, D. A., Salamanca, N., and Zhu, A. (2019). Parenting style as an investment in human development. *Journal of Population Economics*, 32(4):1315–1352.

Cohen-Goldner, S. and Eckstein, Z. (2008). Labor mobility of immigrants: Training, experience, language, and opportunities. *International Economic Review*, 49(3):837–872.

Cohen-Goldner, S. and Eckstein, Z. (2010). Estimating the return to training and occupational experience: The case of female immigrants. *Journal of Econometrics*, 156(1):86–105.

Compton, J. and Pollak, R. A. (2007). Why Are Power Couples Increasingly Concentrated in Large Metropolitan Areas? *Journal of Labor Economics*, 25(3):475–512.

Conti, G., Mason, G., and Poupakis, S. (2019). Developmental Origins of Health Inequality.

Costa, D. L. and Kahn, M. E. (2000). Power Couples: Changes in the Locational Choice of the College Educated, 1940-1990. *The Quarterly Journal of Economics*, 115(4):1287–1315.

Cunha, F. (2015). Subjective Rationality, Parenting Styles, and Investments in Children. In *Families in an Era of Increasing Inequality: Diverging Destinies*, volume 5 of *National Symposium on Family Issues*, pages 83–94. Springer International Publishing, Cham.

Cunha, F., Heckman, J., Lochner, L., and Masterov, D. V. (2006). Interpreting the Evidence on Life Cycle Skill Formation. In *Handbook of the Economics of Education*, pages 697–812. Elsevier.

Cunha, F. and Heckman, J. J. (2008). Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation. *Journal of Human Resources*, 43(4):738–782.

Cunha, F., Heckman, J. J., and Schennach, S. (2010). Estimating the Technology of Cognitive and Noncognitive Skill Formation. *Econometrica*, 78(3):883–931.

Cutler, D. M., Glaeser, E. L., and Vigdor, J. L. (2008). Is the melting pot still hot? Explaining the resurgence of immigrant segregation. *Review of Economics and Statistics*, 90(3):478–497.

Danzer, A. M. and Yaman, F. (2013). Do ethnic enclaves impede immigrants' integration? Evidence from a quasi-experimental social-interaction approach. *Review of International Economics*, 21(2):311–325.

Del Boca, D., Flinn, C., and Wiswall, M. (2014). Household Choices and Child Development. *The Review of Economic Studies*, 81(1):137–185.

Del Boca, D., Flinn, C., and Wiswall, M. (2016). Transfers to Households with Children and Child Development. *The Economic Journal*, 126(596):F136–F183.

Del Bono, E., Francesconi, M., Kelly, Y., and Sacker, A. (2016). Early Maternal Time Investment and Early Child Outcomes. *The Economic Journal*, 126(596):F96–F135.

Delvecchio, E., Germani, A., Raspa, V., Lis, A., and Mazzeschi, C. (2020). Parenting styles and child's well-being: The mediating role of the perceived parental stress. *Europe's Journal of Psychology*, 16(3):514–531.

Deming, D. J. (2017). The Growing Importance of Social Skills in the Labor Market*. *The Quarterly Journal of Economics*, 132(4):1593–1640.

Diamond, R. (2016). The determinants and welfare implications of US Workers' diverging location choices by skill: 1980-2000. *American Economic Review*, 106(3):479–524.

Doepke, M., Sorrenti, G., and Zilibotti, F. (2019). The Economics of Parenting. *Annual Review of Economics*, 11:55–84.

Doepke, M. and Zilibotti, F. (2017). Parenting With Style: Altruism and Paternalism in Intergenerational Preference Transmission. *Econometrica*, 85(5):1331–1371.

Doepke, M. and Zilibotti, F. (2019). The Economics of Parenting Style. In *Love, Money, and Parenting*, pages 21–50. Princeton University Press.

Dooley, M. and Stewart, J. (2007). Family income, parenting styles and child behavioural-emotional outcomes. *Health Economics*, 16(2):145–162.

Dribe, M. and Lundh, C. (2011). Cultural dissimilarity and intermarriage. A longitudinal study of immigrants in Sweden 1990-2005. *International Migration Review*, 45(2):297–324.

Edin, P.-A., Fredriksson, P., Nybom, M., and Öckert, B. (2022). The Rising Return to Noncognitive Skill. *American Economic Journal: Applied Economics*, 14(2):78–100.

Edlund, L. (2006). The Price of Marriage: Net vs. Gross Flows and the South Asian Dowry Debate. *Journal of the European Economic Association*, 4(2-3):542–551.

Elwert, A. (2018). *Will You Intermarry Me? : Determinants and Consequences of Immigrant-Native Intermarriage in Contemporary Nordic Settings.*

Elwert, A. and Tegunimataka, A. (2016). Cohabitation Premiums in Denmark: Income Effects in Immigrant–Native Partnerships. *European Sociological Review*, 32(3):383–402.

Falk, A., Kosse, F., Pinger, P., Schildberg-Hörisch, H., and Deckers, T. (2021). Socioeconomic Status and Inequalities in Children's IQ and Economic Preferences. *Journal of Political Economy*, 129(9):2504–2545.

Fan, J. and Zou, B. (2021). The Dual Local Markets: Family, Jobs, and the Spatial Distribution of Skills. *SSRN Electronic Journal*, (November).

Fiorini, M. and Keane, M. P. (2014). How the Allocation of Children's Time Affects Cognitive and Noncognitive Development. *Journal of Labor Economics*, 32(4):787–836.

Fletcher, A. C., Walls, J. K., Cook, E. C., Madison, K. J., and Bridges, T. H. (2008). Parenting Style as a Moderator of Associations Between Maternal Disciplinary Strategies and Child Well-Being. *Journal of Family Issues*, 29(12):1724–1744.

Foster, A. D. and Rosenzweig, M. R. (2001). Missing women, the marriage market and economic growth.

Fulford, S. (2013). The Puzzle of Marriage Migration in India.

Fulford, S. (2015). Marriage migration in India: Vast, varied, and misunderstood.

Galichon, A. and Salanié, B. (2021). Cupid's Invisible Hand: Social Surplus and Identification in Matching Models. *The Review of Economic Studies*, (December):1–30.

García, F. and Gracia, E. (2009). Is always authoritative the optimum parenting style? Evidence from Spanish families. *Adolescence*, 44(173):101–131.

García-Miralles, E. and Gensowski, M. (2023). Are Children's Socio-Emotional Skills Shaped by Parental Health Shocks? *Journal of Human Resources*, pages 0820–11091R2.

Goodman, A., Lamping, D. L., and Ploubidis, G. B. (2010). When to Use Broader Internalising and Externalising Subscales Instead of the Hypothesised Five Subscales on the Strengths and Difficulties Questionnaire (SDQ): Data from British Parents, Teachers and Children. *Journal of Abnormal Child Psychology*, 38(8):1179–1191.

Grossbard, S. A. and Vernon, V. (2020). Do immigrants pay a price when marrying natives? Lessons from the US time use survey. *IZA Journal of Development and Migration*, 11(1).

Guilmoto, C. Z. and Attané, I. (2007). The Geography of Deteriorating Child Sex Ratio in China and India. In Guilmoto, C. Z. and Attané, I., editors, *Watering the Neighbour's Garden: The Growing Demographic Female Deficit in Asia*, pages 109–129. Committee for International Cooperation in National Research in Demography, Paris.

Guy, S., Furber, G., Leach, M., and Segal, L. (2016). How many children in Australia are at risk of adult mental illness? *Australian & New Zealand Journal of Psychiatry*, 50(12):1146–1160.

Hanushek, E. A. and Woessmann, L. (2008). The role of cognitive skills in economic development. *Journal of Economic Literature*, 46(3):607–668.

Harris, D. R. and Ono, H. (2005). How many interracial marriages would there be if all groups were of equal size in all places? A new look at national estimates of interracial marriage. *Social Science Research*, 34(1):236–251.

Hayfron, J. E. (2001). Language training, language proficiency and earnings of immigrants in Norway. *Applied Economics*, 33(15):1971–1979.

Heberle, A. E., Briggs-Gowan, M. J., and Carter, A. S. (2015). A Person-oriented Approach to Identifying Parenting Styles in Mothers of Early School-age Children: Person-oriented Approach to Parenting Styles. *Infant and Child Development*, 24(2):130–156.

Heckman, J. J. (1979). Sample Selection Bias as a Specification Error. *Econometrica*, 47(1):153–161.

Heckman, J. J. and Mosso, S. (2014). The Economics of Human Development and Social Mobility. *Annual Review of Economics*, 6(1):689–733.

Howenstein, J., Kumar, A., Casamassimo, P. S., McTigue, D., Coury, D., and Yin, H. (2015). Correlating parenting styles with child behavior and caries. *Pediatric Dentistry*, 37(1):59–64.

Hutchison, L., Feder, M., Abar, B., and Winsler, A. (2016). Relations between Parenting Stress, Parenting Style, and Child Executive Functioning for Children with ADHD or Autism. *Journal of Child and Family Studies*, 25(12):3644–3656.

Jayachandran, S. (2015). The Roots of Gender Inequality in Developing Countries. *Annual Review of Economics*, 7(1):63–88.

Jejeebhoy, S. J. and Halli, S. S. (2005). Marriage patterns in rural India: Influence of sociocultural context. *The changing transitions to adulthood in developing countries: Selected studies*, pages 172–199.

Joona, P. A. and Nekby, L. (2012). Intensive Coaching of New Immigrants: An Evaluation Based on Random Program Assignment. *Scandinavian Journal of Economics*, 114(2):575–600.

Kalmijn, M. and van Tubergen, F. (2006). Ethnic intermarriage in the Netherlands: confirmations and refutations of accepted insights. *European Journal of Population / Revue européenne de Démographie*, 22(4):371–397.

Kantarevic, J. (2004). Interethnic Marriages and Economic Assimilation of Immigrants. *IZA Discussion Paper*, (1142).

Kaur, R. (2008). Dispensable Daughters and Bachelor Sons: Sex Discrimination in North India. *Economic and Political Weekly*, 43(30):109–114.

Kaur, R. (2012). Marriage and Migration: Citizenship and Marital Experience in Cross-border Marriages between Uttar Pradesh, West Bengal and Bangladesh. *Economic & Political Weekly*, XLVII(43):78–89.

Kaur, R. (2013). Mapping the Adverse Consequences of Sex Selection and Gender Imbalance in India and China. *Weekly, Economic and Political*, 48(35):37–44.

Kim, J. H. (2019). The Economics of Parenting Skill and Child Development. Technical report.

Klasen, S. and Lahoti, R. (2021). How Serious is the Neglect of Intra-Household Inequality in Multidimensional Poverty and Inequality Analyses? Evidence from India. *Review of Income and Wealth*, 67(3):705–731.

Kukreja, R. and Kumar, P. (2013). Tied in a Knot. Cross-region Marriages in Haryana and Rajasthan: Implications for Gender Rights and Gender Relations. Technical report, Kingston.

Lazear, E. P. (1999). Culture and language. In *Journal of Political Economy*, volume 107, pages S95–S126.

Le Forner, H. (2021). Formation of Children's Cognitive and Socio-Emotional Skills: Is All Parental Time Equal? Technical report.

Lee, S. Y. T. and Seshadri, A. (2019). On the Intergenerational Transmission of Economic Status. *Journal of Political Economy*, 127(2):855–921.

Lehmer, F. and Ludsteck, J. (2015). Wage Assimilation of Foreigners: Which Factors Close the Gap? Evidence From Germany. *Review of Income and Wealth*, 61(4):677–701.

Lise, J. and Seitz, S. (2011). Consumption inequality and intra-household allocations. *Review of Economic Studies*, 78(1):328–355.

Lizzeri, A. and Siniscalchi, M. (2008). Parental Guidance and Supervised Learning. *Quarterly Journal of Economics*, 123(3):1161–1195.

Llull, J. (2018). Immigration, wages, and education: A labour market equilibrium structural model. *Review of Economic Studies*, 85(3):1852–1896.

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

Luyckx, K., Tildesley, E. A., Soenens, B., Andrews, J. A., Hampson, S. E., Peterson, M., and Duriez, B. (2011). Parenting and Trajectories of Children's Maladaptive Behaviors: A 12-Year Prospective Community Study. *Journal of Clinical Child & Adolescent Psychology*, 40(3):468–478.

Maccoby, E. E. and Martin, J. A. (1983). Socialization in the Context of the Family: Parent-Child Interaction. In *In P. H. Mussen, & E. M. Hetherington (Eds.), Handbook of Child Psychology: Vol. 4. Socialization, Personality, and Social Development*, pages 1–101. Wiley, New York.

McWhirter, A. C., McIntyre, L. L., Kosty, D. B., and Stormshak, E. (2023). Parenting Styles, Family Characteristics, and Teacher-Reported Behavioral Outcomes in Kindergarten. *Journal of Child and Family Studies*, 32(3):678–690.

Meng, X. and Gregory, R. G. (2005). Intermarriage and the economic assimilation of immigrants. *Journal of Labor Economics*, 23(1):135–175.

Meng, X. and Meurs, D. (2009). Intermarriage, language, and economic assimilation process: A case study of France. *International Journal of Manpower*, 30(1-2):127–144.

Min Zhou and Logan, J. R. (1989). Returns on human capital in ethnic enclaves: New York City's Chinatown. *American Sociological Review*, 54(5):809–820.

Mishra, P. (2021). When South Meets North: Interrogating Agency and Marital Mobility in Kerala-Haryana Marriages. In *India Studies in Business and Economics*, pages 369–386.

OECD (2020). *International Migration Outlook 2020*. OECD.

Psacharopoulos, G. and Patrinos, H. A. (2004). Returns to investment in education: A further update. *Education Economics*, 12(2):111–134.

Rao, S. and Finnoff, K. (2015). Marriage Migration and Inequality in India, 1983-2008. *Population and Development Review*, 41(3):485–505.

Rao, V. (1993). Dowry 'Inflation' in Rural India: A Statistical Investigation. *Population Studies*, 47(2):283–293.

Roback, J. (1982). Wages , Rents , and the Quality of Life Jennifer Roback. *Journal of Political Economy*, 90(6):1257–1278.

Rodríguez-García, D. (2015). Intermarriage and Integration Revisited: International Experiences and Cross-Disciplinary Approaches. *Annals of the American Academy of Political and Social Science*, 662(1):8–36.

Rosen, K. T. (1979). A Regional Model of Multifamily Housing Starts.

Rosenzweig, M. R. and Stark, O. (1989). Consumption Smoothing, Migration, and Marriage: Evidence from Rural India. *Journal of Political Economy*, 97(4):905–926.

Salem, R. (2018). Matrimonial Expenditures and Egyptian Women's Power Within Marriage. *Journal of Family Issues*, 39(9):2615–2638.

Sanders, M. R. and Woolley, M. L. (2005). The relationship between maternal self-efficacy and parenting practices: Implications for parent training: Self-efficacy and parenting practices. *Child: Care, Health and Development*, 31(1):65–73.

Sautmann, A. (2011). Partner Search and Demographics: The Marriage Squeeze in India. *SSRN Electronic Journal*, 2011(December 2009).

Spera, C. (2005). A Review of the Relationship Among Parenting Practices, Parenting Styles, and Adolescent School Achievement. *Educational Psychology Review*, 17(2):125–146.

Srinivasan, K. and James, K. S. (2015). The Golden Cage: Stability of the Institution of Marriage in India. *Economic and Political Weekly*, 50(13):38–45.

Todd, P. E. and Wolpin, K. I. (2003). On the Specification and Estimation of the Production Function for Cognitive Achievement. *The Economic Journal*, 113(485):F3–F33.

Todd, P. E. and Wolpin, K. I. (2007). The Production of Cognitive Achievement in Children: Home, School, and Racial Test Score Gaps. *Journal of Human Capital*, 1(1):91–136.

Twizeyemariya, A., Guy, S., Furber, G., and Segal, L. (2017). Risks for Mental Illness in Indigenous Australian Children: A Descriptive Study Demonstrating High Levels of Vulnerability: Risks for Mental Illness in Indigenous Australian Children. *The Milbank Quarterly*, 95(2):319–357.

van Tubergen, F. and Maas, I. (2007). Ethnic intermarriage among immigrants in the Netherlands: An analysis of population data. *Social Science Research*, 36(3):1065–1086.

Wiswall, F. and Agostinelli, M. (2020). Estimating the technology of children's skill formation. *NBER Working Paper Series*.

Xie, Y. and Gough, M. (2011). Ethnic Enclaves and the Earnings of Immigrants. *Demography*, 48(4):1293–1315.

A. APPENDIX TO CHAPTER 1

A.1. Descriptive statistics

Table A.1 reports the descriptive statistics for a subsample of immigrants by gender and three marriage statuses: single, inter- and noninter- married. The mean age and standard deviation for intermarried and nonintermarried immigrants are similar for males. The sample of intermarried females seems to be slightly older than nonintermarried ones. Intermarried male and female immigrants, on average, spend more time in Germany by around half a year compared to male immigrants who are not married to German women. Married male immigrants, on average, migrate at the age of 19. In the case of women, intermarried immigrants arrive in the host country at 20, two years older than nonintermarried ones. The difference in years of education between intermarried and nonintermarried immigrants is the same for males and women and equals one year.

The second part of Table A.1 presents statistics associated with immigrants' assimilation. On average, female immigrants married to Germans declare that they feel more German and less often that they do not belong to German society than nonintermarried immigrants. For male immigrants, the relationship is the same regarding feeling that they do not belong to German society. Intermarried immigrants are also characterised by better, on average, knowledge of oral German, and they more often use German media. Immigrants married to other immigrants report being visited by German less frequently than intermarried ones. This data suggest that intermarried immigrants are, on average, better assimilated than nonintermarried ones.

A.2. Data sources

Table A.2: Amenity Indices - data sources

Variable	Source	Sample	Notes
Population	Polizeiliche Kriminalstatistik 1966, 1976, 1986, 1996, 206, 2016	All Bundelands	
Number of passengers in public transport	Statistisches Jahrbuch für die Bundesrepublik Deutschland 1965, 1976, 1986, 1995; Destatis, Tabelle 46181-0011	All public companies	Total number of passengers transported by public companies within calendar year

Length of highways	Statistisches Jahrbuch für die Bundesrepublik Deutschland 1966, 1976, 1986, 1997, 2005; Statistisches Jahrbuch 2015	All public roads	Total length of highways available at the end of the calendar year
Number of national parks	Bundesamt für Naturschutz, Naturparke in Deutschland (01/01/2020)	All national and natural parks	If national parks is a part of more than one Bundesland, then it was assigned to all of them.
Forest area in %	Fachserie B. Land- und Forstwirtschaft, Fischerei. Statistisches Bundesamt Wiesbaden. 1964, 1974, 1985, 1993; Tabelle 33111-0004: Bodenfläche (tatsächliche Nutzung): Bundesländer, Stichtag (bis 31.12.2015), Nutzungsarten	Area classified as a forest	The raw area of forest was recalculated to % using data on area of Bundeslands
Number of crime cases per capita	Sensch, Jürgen (1955-2003 [2005]), histat-Datenkompilation online: Kriminalitätsentwicklung in der Bundesrepublik Deutschland von 1955 bis 2003: Ausgewählte Indikatoren aus der Kriminalstatistik; Polizeiliche Kriminalstatistik 2005, 2015.	Number of recorded cases	
Number of sever crime cases	Polizeiliche Kriminalstatistik 1974, 1978, 1982, 1986, 1990, 1994, 1998, 2002, 2006, 2010, 2014, 2018	Number of murders	
GDP per capita	Statistisches Landesamt Baden-Württemberg, Arbeitskreis "Volkswirtschaftliche Gesamtrechnungen der Länder", "Bruttoinlandsprodukt, Bruttowertschöpfung in den Ländern der Bundesrepublik Deutschland" 1961-2020	All Bundeslands	

Employment per capita	Erwerbstätige in den alten Ländern der Bundesrepublik Deutschland 1970 bis 1991 sowie in deren kreisfreien Städten und Landkreisen 1980, 1985, 1987 bis 1991, Destatis 2005; Erwerbstätige am Arbeitsort Ländererergebnisse – Jahresdurchschnitt, Destatis 1991-2020	All Bundeslands

A.3. The solution of single and couple households' problems at the Stage 3

A.3.1. Couple's problem

This section provides the solution for stage 3 for married couples deciding on their private and public consumption and labor supply.

Individual utilities of agents in couple (H, H^*) at Stage 3 are:

$$u(Q, C, L) = \ln Q(C + \alpha \cdot \ell_{nw}) \tag{A.1}$$

$$u(Q, C^*, L^*) = \ln Q(C^* + \alpha^* \cdot (\ell^*_{pt} + \ell^*_{nw}) + \delta^* \cdot \ell^*_{pt}) \tag{A.2}$$

Preferences satisfy the *transferable utility* (TU) if there exists cardinalisation of representing them utilities, such that far all values of prices and income, the Pareto frontier is a straight line with a slope equal to -1 (Chiappori and Gugl, 2020).

Proposition 3. *Preferences represented by the utility given in Equations A.1 and A.2 satisfy TU property.*

Proof: Assume that we take cardinal representation of the preferences equal to $\exp u_i$ and $\exp u^*$. Then the couple maximisation problem can be written in the following form:

$$\max_{Q,C,C^*} \exp u + \mu \exp u^* \tag{A.3}$$

under the budget constraint:

$$\overline{Y}^{H,H^*}(L, L^*) \geq C + C^* + pQ \tag{A.4}$$

where μ is a Pareto weight and Y represents income of a couple (H, H^*) as a function on their leisure choices. Then using Lagrangia function, the problem can be expressed as:

$$\begin{aligned} \mathcal{L}(Q, C, C^*, \lambda) = Q(C + C^* + \alpha \cdot \ell_{nw} + \alpha^* \cdot (\ell^*_{pt} + \ell^*_{nw}) + \delta^* \cdot \ell^*_{pt}) \\ + \lambda(\overline{Y}^{H,H^*}(L, L^*) - C - C^* - pQ) \end{aligned} \tag{A.5}$$

Table A.1: Descriptive Statistics of Immigrants Subsample

	Male			Female		
	Single	Inter- -married	Noninter -married	Single	Inter- -married	Noninter -married
Age	38.08	39.06	39.83	41.97	37.67	36.72
	(8.71)	(8.75)	(8.64)	(10.18)	(8.01)	(7.83)
Years since mig.	16.57	19.76	19.94	18.89	17.03	18.19
	(11.39)	(11.58)	(10.29)	(12.29)	(10.11)	(9.69)
Mig. age	21.41	19.00	19.55	23.22	20.26	18.32
	(12.62)	(10.09)	(8.85)	(11.97)	(10.03)	(8.48)
Education	11.41	11.35	10.54	11.14	11.55	10.26
	(2.26)	(2.46)	(2.36)	(2.65)	(2.47)	(2.23)
Feel German	2.69	2.86	3.05	2.78	2.69	3.25
	(1.28)	(1.21)	(1.30)	(1.31)	(1.25)	(1.31)
Feel of not belong.	3.91	3.65	3.50	3.30	3.76	3.37
	(0.97)	(1.17)	(1.21)	(0.98)	(1.14)	(1.19)
Oral German skills	2.14	1.83	2.26	1.90	1.85	2.31
	(1.00)	(0.91)	(0.88)	(0.95)	(0.89)	(1.03)
Language of Media	3.66	3.73	3.51	3.89	4.13	3.69
	(1.29)	(1.28)	(1.27)	(1.24)	(1.22)	(1.37)
Visit from German	0.79	0.92	0.83	0.89	0.94	0.83
	(0.41)	(0.27)	(0.37)	(0.31)	(0.24)	(0.38)
Share	16.55%	50.68%	32.76%	13.81%	56.92%	29.27%

Notes: The table lists the mean for demographic characteristics calculated with sample population weights. Standard deviation in the parentheses. Variable Feel German is measured on a 1-5 scale, where 1 means "Completely" and 5 means "Not at All". Variable Feel of not Belonging is measured on a 1-5 scale, where 1 means "Very often" and 5 means "Never". Variable Oral German skills is measured on a 1-5 scale, where 1 means "Very Good" and 5 means "Not at All". Variable Language of Media is measured on a scale 1-5 where 1 means only language country of origin and 5 means only German. Variable Received Visits of Germans is an indicator variable equal 1 if the agent received at least one visit from a German in the previous year.

Taking the derivatives with respect to private consumptions yields:

$$\begin{cases} \frac{\partial L}{\partial C} & = Q - \lambda = 0 \\ \frac{\partial L}{\partial C^*} & = \mu Q - \lambda = 0 \end{cases} \implies \mu = 1 \tag{A.6}$$

Transferable utility implies that household aggregate demand does not depend on Pareto weight μ. So, the household (H, H^*) maximization problem at Stage 3 is as follows:

$$\max_{\overline{C},Q} Q(\overline{C} + \Lambda_m(L) + \Lambda_f(L^*)) \tag{A.7}$$

where

$$\Lambda(L)_g \equiv \begin{cases} \alpha\ell_{nw} & \text{if } g = m \\ \alpha\ell_{nw} + \delta\ell_{pt} & \text{otherwise} \end{cases} \tag{A.8}$$

with respect to the budget constraint:

$$\overline{Y}^{H,H^*}(L, L^*) \equiv y_{nl}(H, H^*) + \ell_{nw} \cdot b(w) + \ell^*_{nw} \cdot b(w^*) + w_{net}(L, L^*, w, w^*) \tag{A.9}$$

$$= \overline{C} + pQ. \tag{A.10}$$

Conditioning on labor supply (L, L^*) the ex-post (after realization of productivity and leisure preference shocks) efficient allocation is as follows:

$$pQ(L, L^*) = (\overline{Y}^{H,H^*}(L, L^*) + \Lambda_m(L) + \Lambda_f(L^*))/2 \tag{A.11}$$

$$\overline{C}(L, L^*) = (\overline{Y}^{H,H^*}(L, L^*) - \Lambda_m(L) - \Lambda_f(L^*))/2 = pQ - \Lambda_m(L) - \Lambda_f(L^*) \tag{A.12}$$

The equations A.11 and A.12 describe the aggregated demand of couple (H, H^*) for private and public consumptions as a function of individual labor supply choices. Using this fact, the optimal labor supply can be found by solving the following maximization problem:

$$\max_{L,L^*} pQ^2(L, L^*). \tag{A.13}$$

The final maximization problem is a discrete choice problem. Each couple (H, H^*) has 3×2 possible labor supply choices. Given the solution to this problem, one can recover aggregated demands for private and public consumptions of union (H, H^*).

Let's define $C^*(L, L^*) = (\overline{Y}^{H,H^*}(L, L^*) - \Lambda_m(L) - \Lambda_f(L^*))/2 - C(L, L^*)$. Ex-ante (at Stage 2, before realization of productivity and leisure preference shocks) efficiency requires that C maximizes some weighted sum of individual expected utilities, formally:

$$\max_{C} \mathrm{E}\, u + \mu\, \mathrm{E}\, u^* \tag{A.14}$$

for some $\mu > 0$, under the resources constraint given by A.9. First-order condition implies:

$$\frac{\partial}{\partial C} = \frac{1}{C + \Lambda_m(L)} - \frac{\mu}{(\overline{Y}^{H,H^*}(L, L^*) - \Lambda_m(L) - \Lambda_f(L^*))/2 - C - \Lambda_f(L^*)} = 0 \tag{A.15}$$

As a result, private consumption of agents is given by:

$$C = \frac{\overline{Y}^{H,H^*}(L,L^*) + \Lambda_m(L) + \Lambda_f(L^*)}{2(\mu+1)} - \Lambda(L) = \frac{1}{1+\mu} pQ(L,L^*) - \Lambda_m(L) \tag{A.16}$$

$$C^* = \mu \cdot \frac{\overline{Y}^{H,H^*}(L,L^*) + \Lambda_m(L) + \Lambda_f(L^*)}{2(\mu+1)} - \Lambda(L^*) = \frac{\mu}{1+\mu} pQ(L,L^*) - \Lambda_f(L^*) \tag{A.17}$$

$$\tag{A.18}$$

Finally, individual expected utilities are equal to the following:

$$\mathrm{E}\,u = \ln p + \ln\frac{1}{1+\mu} + \int \ln Q^2(H,H^*,r_{ij},\boldsymbol{\varepsilon},\boldsymbol{\upsilon},\boldsymbol{\zeta})dF(\boldsymbol{\varepsilon},\boldsymbol{\upsilon},\boldsymbol{\zeta}) \tag{A.19}$$

$$\mathrm{E}\,u^* = \ln p + \ln\frac{\mu}{1+\mu} + \int \ln Q^2(H,H^*,r_{ij},\boldsymbol{\varepsilon},\boldsymbol{\upsilon},\boldsymbol{\zeta})dF(\boldsymbol{\varepsilon},\boldsymbol{\upsilon},\boldsymbol{\zeta}) \tag{A.20}$$

where F denotes the joint distribution of productivity and leisure shocks.

Let $\Psi(H,H^*,r)$ denotes the common part of private consumption:

$$\Psi(H,H^*,r) = \ln p + \int \ln Q^2(H,H^*,r_{ij},\boldsymbol{\varepsilon},\boldsymbol{\upsilon},\boldsymbol{\zeta})dF(\boldsymbol{\varepsilon},\boldsymbol{\upsilon},\boldsymbol{\zeta}). \tag{A.21}$$

Then, taking exp of both sides and adding up gives the set of ex-ante (at Stage 2) Pareto efficient allocations:

$$\exp\{\mathrm{E}\,u\} + \exp\{\mathrm{E}\,u^*\} = \frac{1}{1+\mu}\exp\{\Psi(H,H^*,r)\} + \frac{\mu}{1+\mu}\exp\{\Psi(H,H^*,r)\} \tag{A.22}$$

$$= \exp\{\Psi(H,H^*,r)\} = \overline{U}(H,H^*,r) \tag{A.23}$$

which is a TU form for $U_g^H(H^*,r) = \exp\{\mathrm{E}\,u\}$ and $U_{g^*}^{H^*}(H,r) = \exp\{\mathrm{E}\,u^*\}$.

A.3.2. Single's problem

This section provides the solution for stage 3 for singles who choose their private and public consumption and labor supply. Single agents at stage 3 face the following maximization problem:

At the stage 3 of the model single household $(H,\varnothing)$ (equivalently for $(\varnothing,H_j)$) solves the following maximization problem:

$$\max_{C,Q} Q(C + \Lambda_g(L)) \tag{A.24}$$

with respect to the budget constrain:

$$Y^H(L) \equiv y_{nl}(H) + \ell_{nw}\cdot b(w) + w_{net}(w,L) = C + pQ. \tag{A.25}$$

where $\Lambda_g(L)$ is defined as in Equation A.8.

The maximization problem of single agents is very similar to that of a couple since the union of (H, H^*) at stage 3 behaves as a single decision maker. Then, the conditional on the labor supply choice demand for private and public goods is given by:

$$pQ(L) = (Y^H(L) + \Lambda(L))/2 \tag{A.26}$$

$$C(L) = (Y^H(L) - \Lambda(L))/2 = pQ - \Lambda(L) \tag{A.27}$$

Using Equations A.26 and A.27 to substitute Q and C in the Equation A.24, one gets the expression, which can be used to find the optimal labor supply. Single i finds the optimal labor supply by solving the following maximization problem:

$$\max_L pQ^2(L) \tag{A.28}$$

It is a discrete choice problem, where every single agent has three (or two for men) possible choices. Given the solution to this problem, one can recover demand for the single agent's public and private consumption.

Then expected utility (at Stage 2, before realization of productivity and leisure preference shocks) is given by:

$$\mathrm{E}\, u = \ln p + \int \ln Q^2(H, r, \varepsilon, \upsilon, \zeta) dF(\varepsilon, \upsilon, \zeta)$$

Finally define the exponential representation of the utility function $U_g^H(\varnothing, r) \equiv \exp\{\mathrm{E}\, u\}$, which corresponds to a TU form from the marriage problem.

A.4. Identification of marriage market parameters and sharing rule

Let N_r^H be a number of men with human capital H who live in region r. Then, $N_{d,r}^{H,H^*}$ is the number of (H, H^*) marriages demanded by men with human capital H and $N_{d,r}^{H,\varnothing}$ is the number of unmarried men with human capital H. Using Equation 1.20 and the fact that ML estimator of $\mathrm{P}(H^*|H, r)$ is $\frac{N_{d,r}^{H,H^*}}{N_r^H}$, I derive a quasi-demand equation for men:

$$\begin{aligned} \ln \frac{N_{d,r}^{H,H^*}}{N_{d,r}^{H,\varnothing}} &= \Gamma_M(H, H^{\cdot} r*) - \Gamma_M(H, \varnothing, r) = \\ &= (\overline{U}(H, H^*, r) - \tau(H, H^*, r) - U_g(H, \varnothing, r) + \phi_1|e^* - e| + \phi_2|o^* - o| - \phi_{0H})/\sigma_\omega^{M,H} \end{aligned} \tag{A.29}$$

and a quasi-supply equation for women:

$$\begin{aligned} \ln \frac{N_{s,r}^{H,H^*}}{N_{s,r}^{\varnothing,H^*}} &= \Gamma_F(H^*, H, r) - \Gamma_F(H^*, \varnothing, r) \\ &= (\tau(H, H^*, r) - U_{g^*}(H^*, \varnothing, r) + \phi_1|e - e^*| + \phi_2|o - o^*| - \phi_{0H^*})/\sigma_\omega^{F,H^*} \end{aligned} \tag{A.30}$$

In each location, r 4×4 sub-marriage market clears, when given equilibrium transfers τ's, the demand by men with H for women with H^* is equal to the supply of women with H^* for men with H for all possible combinations of the human capital. Finally, the identification of transfers τ's and taste for similarity phi's can be obtained using Equations A.29 and A.30.

To identify μ's so Pareto weights associated with the initial maximization problem, I use Equation A.22 and show that:

$$\tau(H, H^*, r) = U_{g^*}(H^*, H, r) = \frac{\mu(H, H^*, r)}{1 + \mu(H, H^*, r)} \times \overline{U}(H, H^*, r) \tag{A.31}$$

Solving for $\mu(H, H^*, r)$ yields:

$$\mu(H, H^*, r) = \frac{\tau(H, H^*, r)}{\overline{U}(H, H^*, r) - \tau(H, H^*, r)} \tag{A.32}$$

$\mu(H, H^*, r)$ is well-defined if $\tau(H, H^*, r) \in (0, \overline{U}(H, H^*, r))$. So, if $\tau(H, H^*, r)$ is identified, then the $\mu(H, H^*, r)$ is also identified.

A.5. German social security and tax systems

A.5.1. German tax code

In Germany, each employee pays two types of social contribution: social system contribution and personal income tax. Social system contribution depends on individual yearly labor income. However, married couples in Germany submit tax statements together. As a result, the amount of paid personal income tax depends on the yearly labor income of both partners.

To approximate the level of social contribution, first, I define yearly individual labor income as a function of wage and labor supply choices. I assume that every full-time employed agent works approximately 1778 hours per year, while the part-time employed agent works half of it. So, the individual yearly labor income y is obtained in the following way:

$$y(w, L) = w \cdot [(\ell_{ft} + 0.5\ell_{pt}) \cdot 1778]$$

In Germany, individuals who earn less than 4.800€ do not pay social system contributions. There is also a maximum amount to contribute to the social system. This amount slightly changes every year. Since this paper uses data from 1984 to 2018, I take the threshold from 2005 (13104€) as a representative for the whole sample.

Then, I use the following piece-wise function of yearly individual labor income to approximate the share of gross income which contributes to the social system:

$$\tau_{sc}(y) = \begin{cases} 0 & \text{if } y \leq 4800 \\ 0.0002625 \cdot y & \text{if } 4800 < y \leq 9600 \\ 0.21 & \text{if } 9600 < y \leq 62400 \\ 13104/y & \text{otherwise} \end{cases}$$

The income tax rate is calculated using taxable income, which I take as income after the social security contribution. For agent with H it is $y_{sc}^{H} = (1 - \tau_{sc}(y^{H})) \cdot y^{H}$. As I mentioned, married couples are taxed jointly in Germany as if each earned half of the joint income. This situation can be especially beneficial when there is a big gap between partners' incomes. It provides incentives for one of the partners to work less.

Assume that y_{sc}^{H,H^*} is an average taxable couple's yearly labor income. The tax schedule changes slightly in Germany every year. For consistency, I use a tax schedule for 2005 as a representative for my sample. So, the tax rate of individuals in a couple (H, H^*) is approximated in the following way:

$$\tau_{pit}(y_{sc}^{H,H^*}) = \begin{cases} 0 & \text{if } y_{sc}^{H,H^*} \leq 7664 \\ ((883.74 \cdot \hat{y}^{H,H^*} + 1500) \cdot \hat{y}^{H,H^*})/y_{sc}^{H,H^*} & \text{if } 7664 < y_{sc}^{H,H^*} \leq 12740 \\ ((228.74 \cdot \bar{y}^{H,H^*} + 2397) \cdot \bar{y}^{H,H^*} + 989)/y_{sc}^{H,H^*} & \text{if } 12740 < y_{sc}^{H,H^*} < 52152 \\ (0.42 \cdot y_{sc}^{H,H^*} - 7914)/y_{sc}^{H,H^*} & \text{otherwise} \end{cases}$$

where:

$$\hat{y}^{H,H^*} = (y_{sc}^{H,H^*} - 7644)/10000 \qquad \bar{y}^{H,H^*} = (y_{sc}^{H,H^*} - 12740)/10000$$

In case of single agents y_{sc}^{H,H^*} is replaced by individual taxable yearly labor income y_{sc}^{H}.

A.5.2. Unemployment benefit

In Germany, unemployment benefit is a percentage of the last obtained income and is bounded from above by a certain threshold set by the government. The rules determining the size of unemployment benefit change over time. Since in this paper I use data from 1984 to 2017, I take the threshold (8.68€/h) and percentage of the last obtained income (60%) from 2005 as a representative for the whole sample.

Given the static structure of the model, there is no information about the last period's income. To approximate an unemployment benefit, first, I approximate the last obtained income using an expected wage $\mathrm{E}[w]$ (so wage w net of productivity shock) and assume that an agent worked full-time. Then, I calculate the size of an unemployment benefit using the following formula:

$$b(w) = \min(0.6 \cdot w_{net}(\ell_{ft} = 1, E[w]), 8.68) \tag{A.33}$$

A.6. Model fit

This subsection of appendix contains a set of tables showing the fitness of the model. Header "Simulation" refers to moments obtained in the simulation. The data moments are included under header "Data", while their standard errors are presented under the header "Data SE". Finally, header "Diff in SE" corresponds to the difference between simulated and data moments expressed in standard deviations.

Table A.3: Log wage, married, native male

Moment	Simulation	Data	Data SE	Diff in SE
Noncollege				
Mean	2.695	2.693	0.004	0.548
Variance	0.158	0.208	0.052	0.951
P(wage<Q10)	0.101	0.100	0.009	0.103
P(wage<Q25)	0.307	0.250	0.005	11.864
P(wage<Q50)	0.538	0.500	0.004	9.644
P(wage<Q75)	0.748	0.750	0.005	0.403
P(wage<Q90)	0.888	0.900	0.006	1.885
College				
Mean	3.202	3.197	0.008	0.612
Variance	0.246	0.229	0.072	0.229
P(wage<Q10)	0.142	0.100	0.019	2.209
P(wage<Q25)	0.355	0.250	0.008	12.359
P(wage<Q50)	0.540	0.500	0.007	5.468
P(wage<Q75)	0.715	0.750	0.009	4.053
P(wage<Q90)	0.855	0.900	0.013	3.517

Table A.4: Log wage, married, immigrant male

Moment	Simulation	Data	Data SE	Diff in SE
Noncollege				
Mean	2.600	2.583	0.012	1.314
Variance	0.142	0.210	0.089	0.756
P(wage<Q10)	0.101	0.100	0.022	0.043
P(wage<Q25)	0.300	0.250	0.014	3.629
P(wage<Q50)	0.540	0.500	0.011	3.590
P(wage<Q75)	0.741	0.750	0.013	0.724
P(wage<Q90)	0.879	0.900	0.016	1.268
College				
Mean	3.136	3.112	0.035	0.667
Variance	0.194	0.291	0.165	0.588
P(wage<Q10)	0.046	0.100	0.075	0.715
P(wage<Q25)	0.276	0.255	0.052	0.419
P(wage<Q50)	0.585	0.500	0.038	2.233
P(wage<Q75)	0.773	0.750	0.038	0.591
P(wage<Q90)	0.891	0.905	0.045	0.311

Table A.5: Log wage, married, native female

Moment	Simulation	Data	Data SE	Diff in SE
Noncollege				
Mean	2.383	2.378	0.006	0.883
Variance	0.271	0.357	0.070	1.225
P(wage<Q10)	0.072	0.100	0.015	1.850
P(wage<Q25)	0.254	0.250	0.008	0.444
P(wage<Q50)	0.547	0.500	0.006	7.588
P(wage<Q75)	0.763	0.750	0.005	2.468
P(wage<Q90)	0.887	0.900	0.007	1.794
College				
Mean	2.939	2.940	0.013	0.057
Variance	0.210	0.353	0.105	1.367
P(wage<Q10)	0.060	0.100	0.035	1.160
P(wage<Q25)	0.290	0.250	0.017	2.390
P(wage<Q50)	0.563	0.500	0.012	5.015
P(wage<Q75)	0.786	0.750	0.013	2.769
P(wage<Q90)	0.910	0.900	0.018	0.541

Table A.6: Log wage, married, immigrant female

Moment	Simulation	Data	Data SE	Diff in SE
Noncollege				
Mean	2.302	2.289	0.017	0.741
Variance	0.293	0.315	0.117	0.188
P(wage<Q10)	0.115	0.100	0.037	0.410
P(wage<Q25)	0.279	0.250	0.021	1.347
P(wage<Q50)	0.480	0.500	0.019	1.101
P(wage<Q75)	0.731	0.751	0.020	0.963
P(wage<Q90)	0.871	0.900	0.023	1.265
College				
Mean	2.763	2.745	0.039	0.444
Variance	0.239	0.446	0.193	1.073
P(wage<Q10)	0.037	0.101	0.084	0.759
P(wage<Q25)	0.230	0.251	0.057	0.373
P(wage<Q50)	0.519	0.502	0.047	0.371
P(wage<Q75)	0.809	0.750	0.045	1.306
P(wage<Q90)	0.946	0.900	0.050	0.921

Table A.7: Log wage, single, native male

Moment	Simulation	Data	Data SE	Diff in SE
Noncollege				
Mean	2.604	2.553	0.015	3.374
Variance	0.138	0.281	0.106	1.350
P(wage<Q10)	0.036	0.100	0.035	1.858
P(wage<Q25)	0.223	0.250	0.019	1.432
P(wage<Q50)	0.550	0.500	0.015	3.298
P(wage<Q75)	0.782	0.750	0.014	2.195
P(wage<Q90)	0.898	0.900	0.016	0.153
College				
Mean	3.033	3.001	0.024	1.329
Variance	0.224	0.275	0.133	0.385
P(wage<Q10)	0.062	0.100	0.049	0.788
P(wage<Q25)	0.297	0.251	0.031	1.475
P(wage<Q50)	0.548	0.502	0.026	1.785
P(wage<Q75)	0.742	0.750	0.027	0.284
P(wage<Q90)	0.877	0.901	0.036	0.644

Table A.8: Log wage, single, immigrant male

Moment	Simulation	Data	Data SE	Diff in SE
Noncollege				
Mean	2.406	2.406	0.048	0.006
Variance	0.161	0.222	0.178	0.344
P(wage<Q10)	0.087	0.105	0.093	0.191
P(wage<Q25)	0.241	0.251	0.062	0.159
P(wage<Q50)	0.485	0.501	0.055	0.279
P(wage<Q75)	0.768	0.755	0.056	0.222
P(wage<Q90)	0.916	0.908	0.067	0.120
College				
Mean	3.115	3.078	0.090	0.414
Variance	0.182	0.324	0.271	0.526
P(wage<Q10)	0.073	0.103	0.109	0.278
P(wage<Q25)	0.265	0.264	0.104	0.010
P(wage<Q50)	0.557	0.502	0.113	0.490
P(wage<Q75)	0.796	0.756	0.107	0.374
P(wage<Q90)	0.902	0.918	0.099	0.162

Table A.9: Log wage, single, native female

Moment	Simulation	Data	Data SE	Diff in SE
Noncollege				
Mean	2.495	2.461	0.015	2.297
Variance	0.241	0.325	0.109	0.766
P(wage<Q10)	0.078	0.100	0.041	0.540
P(wage<Q25)	0.298	0.250	0.020	2.347
P(wage<Q50)	0.562	0.500	0.013	4.656
P(wage<Q75)	0.731	0.751	0.013	1.586
P(wage<Q90)	0.850	0.900	0.016	3.253
College				
Mean	2.932	2.891	0.022	1.880
Variance	0.218	0.301	0.131	0.633
P(wage<Q10)	0.062	0.100	0.064	0.605
P(wage<Q25)	0.308	0.250	0.029	2.020
P(wage<Q50)	0.564	0.500	0.020	3.178
P(wage<Q75)	0.724	0.750	0.020	1.305
P(wage<Q90)	0.866	0.901	0.027	1.281

Table A.10: Log wage, single, immigrant female

Moment	Simulation	Data	Data SE	Diff in SE
Noncollege				
Mean	2.285	2.299	0.052	0.269
Variance	0.231	0.360	0.210	0.614
P(wage<Q10)	0.093	0.102	0.072	0.125
P(wage<Q25)	0.221	0.260	0.064	0.610
P(wage<Q50)	0.507	0.508	0.063	0.020
P(wage<Q75)	0.784	0.751	0.058	0.576
P(wage<Q90)	0.882	0.901	0.053	0.360
College				
Mean	2.750	2.732	0.082	0.221
Variance	0.202	0.360	0.266	0.594
P(wage<Q10)	0.039	0.105	0.187	0.352
P(wage<Q25)	0.199	0.267	0.115	0.586
P(wage<Q50)	0.544	0.510	0.109	0.308
P(wage<Q75)	0.840	0.755	0.098	0.860
P(wage<Q90)	0.896	0.903	0.079	0.083

Table A.11: Log wage, native male

Educ.	Partner o_j	Moment	Simulation	Data	Data SE	Diff in SE
Noncollege	native	Mean	2.695	2.691	0.004	0.830
		Variance	0.158	0.208	0.053	0.950
	immigrant	Mean	2.700	2.716	0.017	0.905
		Variance	0.159	0.198	0.103	0.380
College	native	Mean	3.198	3.195	0.008	0.340
		Variance	0.245	0.229	0.073	0.222
	immigrant	Mean	3.237	3.215	0.029	0.740
		Variance	0.251	0.234	0.142	0.119

Table A.12: Log wage, immigrant male

Educ.	Partner o_j	Moment	Simulation	Data	Data SE	Diff in SE
Noncollege	native	Mean	2.614	2.591	0.021	1.097
		Variance	0.141	0.257	0.123	0.946
	immigrant	Mean	2.588	2.577	0.014	0.784
		Variance	0.143	0.167	0.090	0.265
College	native	Mean	3.158	3.132	0.041	0.611
		Variance	0.189	0.273	0.175	0.479
	immigrant	Mean	3.112	3.087	0.062	0.414
		Variance	0.198	0.314	0.222	0.526

Table A.13: Log wage, native female

Educ.	Partner o_j	Moment	Simulation	Data	Data SE	Diff in SE
Noncollege	native	Mean	2.382	2.376	0.006	0.876
		Variance	0.272	0.357	0.071	1.206
	immigrant	Mean	2.405	2.403	0.027	0.073
		Variance	0.263	0.354	0.150	0.608
College	native	Mean	2.937	2.938	0.013	0.103
		Variance	0.210	0.348	0.106	1.306
	immigrant	Mean	2.963	2.958	0.055	0.081
		Variance	0.209	0.418	0.225	0.929

Table A.14: Log wage, immigrant female

Educ.	Partner o_j	Moment	Simulation	Data	Data SE	Diff in SE
Noncollege	native	Mean	2.312	2.290	0.027	0.827
		Variance	0.298	0.334	0.148	0.244
	immigrant	Mean	2.291	2.288	0.020	0.168
		Variance	0.287	0.292	0.125	0.037
College	native	Mean	2.743	2.725	0.047	0.386
		Variance	0.243	0.478	0.215	1.093
	immigrant	Mean	2.813	2.804	0.070	0.129
		Variance	0.224	0.354	0.244	0.533

Table A.15: Mean of log wage by region and education,native male

Region	Education	Simulation	Data	Data SE	Diff in SE
North	Noncollege	2.635	2.614	0.009	2.376
North	College	3.102	3.085	0.016	1.088
South	Noncollege	2.711	2.702	0.006	1.574
South	College	3.202	3.204	0.011	0.142
West	Noncollege	2.683	2.684	0.007	0.101
West	College	3.188	3.166	0.013	1.722

Table A.16: Mean of log wage by region and education,immigrant male

Region	Education	Simulation	Data	Data SE	Diff in SE
North	Noncollege	2.224	2.184	0.042	0.956
North	College	2.565	2.529	0.105	0.350
South	Noncollege	2.358	2.339	0.023	0.812
South	College	2.852	2.807	0.048	0.937
West	Noncollege	2.257	2.265	0.028	0.291
West	College	2.713	2.755	0.056	0.755

Table A.17: Mean of log wage by region and education,native female

Region	Education	Simulation	Data	Data SE	Diff in SE
North	Noncollege	0.868	0.869	0.003	0.280
North	College	0.895	0.896	0.005	0.215
South	Noncollege	0.715	0.714	0.004	0.262
South	College	0.831	0.857	0.007	3.566
West	Noncollege	0.711	0.711	0.011	0.006
West	College	0.832	0.826	0.015	0.349

Table A.18: Mean of log wage by region and education,immigrant female

Region	Education	Simulation	Data	Data SE	Diff in SE
North	Noncollege	0.649	0.636	0.032	0.404
North	College	0.797	0.756	0.050	0.827
South	Noncollege	0.388	0.319	0.011	6.156
South	College	0.375	0.344	0.023	1.311
West	Noncollege	0.178	0.223	0.028	1.610
West	College	0.153	0.229	0.049	1.560

Table A.19: Probability of working ($\ell^{ft} + \ell^{pt} = 1$), married agents

		Simulation	Data	Data SE	Diff in SE
Male					
	native x noncollege	0.868	0.869	0.003	0.280
	native x college	0.895	0.896	0.005	0.215
	immig x noncollege	0.730	0.747	0.011	1.564
	immig x college	0.681	0.693	0.026	0.428
Female					
	native x noncollege	0.715	0.714	0.004	0.262
	native x college	0.831	0.857	0.007	3.566
	immig x noncollege	0.584	0.578	0.012	0.496
	immig x college	0.640	0.639	0.023	0.036

Table A.20: Probability of working ($\ell^{ft} + \ell^{pt} = 1$), single agents

		Simulation	Data	Data SE	Diff in SE
Male					
	native x noncollege	0.711	0.711	0.011	0.006
	native x college	0.832	0.826	0.015	0.349
	immig x noncollege	0.622	0.571	0.037	1.373
	immig x college	0.633	0.617	0.061	0.264
Female					
	native x noncollege	0.827	0.827	0.009	0.038
	native x college	0.931	0.936	0.009	0.481
	immig x noncollege	0.649	0.636	0.032	0.404
	immig x college	0.797	0.756	0.050	0.827

Table A.21: Probability of part-time working ($\ell^{pt} = 1$), females

Region	Education	Simulation	Data	Data SE	Diff in SE
Married					
	native x noncollege	0.380	0.378	0.004	0.516
	native x college	0.275	0.338	0.010	6.655
	immig x noncollege	0.388	0.319	0.011	6.156
	immig x college	0.375	0.344	0.023	1.311
Single					
	native x noncollege	0.164	0.175	0.009	1.285
	native x college	0.134	0.150	0.014	1.160
	immig x noncollege	0.178	0.223	0.028	1.610
	immig x college	0.153	0.229	0.049	1.560

Table A.22: Probability of marriage, male born in '50s living in North

	Female Human Capital				
Male Human Capital	Native noncollege	Native college	Immigrant noncollege	Immigrant college	Single
Simulation					
Native noncollege	0.761	0.040	0.030	0.004	0.165
Native college	0.552	0.282	0.011	0.016	0.139
Immigrant noncollege	0.314	0.026	0.461	0.038	0.161
Immigrant college	0.154	0.141	0.134	0.295	0.275
Data					
Native noncollege	0.749	0.043	0.031	0.004	0.174
Native college	0.428	0.366	0.016	0.022	0.168
Immigrant noncollege	0.293	0.024	0.454	0.079	0.150
Immigrant college	0.160	0.201	0.126	0.364	0.149
Diff in SE					
Native noncollege	1.681	0.877	0.155	0.092	1.370
Native college	8.890	6.233	1.363	1.428	2.713
Immigrant noncollege	0.990	0.280	0.271	3.181	0.648
Immigrant college	0.183	1.638	0.276	1.592	3.932

Table A.23: Probability of marriage, male born in '50s living in South

Male Human Capital	Female Human Capital				
	Native noncollege	Native college	Immigrant noncollege	Immigrant college	Single
Simulation					
Native noncollege	0.780	0.042	0.036	0.004	0.138
Native college	0.588	0.335	0.017	0.021	0.038
Immigrant noncollege	0.296	0.024	0.526	0.035	0.118
Immigrant college	0.175	0.169	0.197	0.417	0.041
Data					
Native noncollege	0.800	0.034	0.039	0.004	0.123
Native college	0.515	0.319	0.030	0.027	0.109
Immigrant noncollege	0.273	0.022	0.562	0.057	0.086
Immigrant college	0.179	0.159	0.192	0.370	0.101
Diff in SE					
Native noncollege	4.286	3.647	1.267	0.252	4.000
Native college	7.252	1.739	3.710	2.046	11.137
Immigrant noncollege	1.951	0.615	2.755	3.583	4.429
Immigrant college	0.158	0.484	0.240	1.688	3.412

Table A.24: Probability of marriage, male born in '50s living in West

Male Human Capital	Female Human Capital				
	Native noncollege	Native college	Immigrant noncollege	Immigrant college	Single
Simulation					
Native noncollege	0.745	0.041	0.033	0.004	0.178
Native college	0.541	0.306	0.014	0.017	0.122
Immigrant noncollege	0.284	0.024	0.471	0.033	0.188
Immigrant college	0.158	0.160	0.161	0.311	0.212
Data					
Native noncollege	0.807	0.033	0.033	0.004	0.124
Native college	0.523	0.332	0.021	0.020	0.105
Immigrant noncollege	0.283	0.022	0.562	0.052	0.081
Immigrant college	0.179	0.161	0.180	0.385	0.095
Diff in SE					
Native noncollege	11.959	3.501	0.042	0.099	12.489
Native college	1.518	2.238	1.935	0.859	2.259
Immigrant noncollege	0.064	0.419	5.707	2.675	12.214
Immigrant college	0.726	0.037	0.648	1.990	5.140

Table A.25: Probability of marriage, male born in '60s living in North

Male Human Capital	Female Human Capital Native noncollege	Native college	Immigrant noncollege	Immigrant college	Single
Simulation					
Native noncollege	0.686	0.045	0.031	0.005	0.234
Native college	0.513	0.267	0.014	0.016	0.190
Immigrant noncollege	0.274	0.030	0.454	0.042	0.199
Immigrant college	0.157	0.136	0.145	0.278	0.283
Data					
Native noncollege	0.666	0.047	0.027	0.006	0.255
Native college	0.392	0.346	0.019	0.029	0.214
Immigrant noncollege	0.242	0.030	0.480	0.073	0.176
Immigrant college	0.155	0.138	0.146	0.319	0.243
Diff in SE					
Native noncollege	3.189	0.974	2.116	1.002	3.590
Native college	10.554	7.018	1.555	3.302	2.554
Immigrant noncollege	1.994	0.006	1.350	3.096	1.642
Immigrant college	0.086	0.057	0.025	1.124	1.215

Table A.26: Probability of marriage, male born in '60s living in South

Male Human Capital	Female Human Capital Native noncollege	Native college	Immigrant noncollege	Immigrant college	Single
Simulation					
Native noncollege	0.712	0.047	0.038	0.005	0.199
Native college	0.568	0.330	0.023	0.021	0.058
Immigrant noncollege	0.260	0.029	0.522	0.040	0.149
Immigrant college	0.179	0.163	0.214	0.402	0.042
Data					
Native noncollege	0.733	0.042	0.039	0.006	0.180
Native college	0.469	0.309	0.028	0.033	0.160
Immigrant noncollege	0.235	0.022	0.579	0.053	0.111
Immigrant college	0.200	0.134	0.205	0.320	0.140
Diff in SE					
Native noncollege	4.832	2.567	0.785	1.764	4.970
Native college	12.403	2.839	2.187	4.137	17.439
Immigrant noncollege	2.656	1.956	5.114	2.630	5.426
Immigrant college	1.128	1.789	0.433	3.674	5.907

Table A.27: Probability of marriage, male born in '60s living in West

Male Human Capital	Female Human Capital				
	Native noncollege	Native college	Immigrant noncollege	Immigrant college	Single
Simulation					
Native noncollege	0.668	0.045	0.033	0.004	0.249
Native college	0.505	0.293	0.017	0.017	0.168
Immigrant noncollege	0.248	0.028	0.459	0.037	0.228
Immigrant college	0.160	0.154	0.170	0.298	0.219
Data					
Native noncollege	0.736	0.043	0.035	0.004	0.183
Native college	0.489	0.319	0.021	0.022	0.149
Immigrant noncollege	0.232	0.021	0.572	0.049	0.126
Immigrant college	0.202	0.180	0.198	0.306	0.114
Diff in SE					
Native noncollege	13.364	1.191	0.500	0.162	14.833
Native college	1.569	2.683	1.188	1.795	2.525
Immigrant noncollege	1.432	1.815	8.456	2.092	11.363
Immigrant college	1.619	1.050	1.055	0.265	5.028

Table A.28: Probability of marriage, male born in '70s living in North

Male Human Capital	Female Human Capital				
	Native noncollege	Native college	Immigrant noncollege	Immigrant college	Single
Simulation					
Native noncollege	0.616	0.054	0.057	0.008	0.265
Native college	0.382	0.291	0.026	0.033	0.267
Immigrant noncollege	0.223	0.028	0.487	0.048	0.214
Immigrant college	0.099	0.129	0.141	0.283	0.348
Data					
Native noncollege	0.582	0.061	0.049	0.010	0.298
Native college	0.307	0.357	0.023	0.050	0.264
Immigrant noncollege	0.196	0.025	0.532	0.057	0.190
Immigrant college	0.094	0.133	0.156	0.332	0.286
Diff in SE					
Native noncollege	4.215	1.969	2.234	1.148	4.319
Native college	6.242	5.165	0.724	2.917	0.280
Immigrant noncollege	1.956	0.686	2.615	1.080	1.711
Immigrant college	0.278	0.164	0.571	1.451	1.915

Table A.29: Probability of marriage, male born in '70s living in South

Male Human Capital	Female Human Capital: Native noncollege	Native college	Immigrant noncollege	Immigrant college	Single
Simulation					
Native noncollege	0.640	0.057	0.069	0.008	0.227
Native college	0.440	0.376	0.042	0.049	0.093
Immigrant noncollege	0.211	0.027	0.558	0.046	0.158
Immigrant college	0.119	0.163	0.218	0.443	0.058
Data					
Native noncollege	0.626	0.064	0.074	0.013	0.222
Native college	0.337	0.341	0.042	0.058	0.222
Immigrant noncollege	0.243	0.023	0.551	0.056	0.128
Immigrant college	0.113	0.134	0.185	0.361	0.206
Diff in SE					
Native noncollege	2.211	2.092	1.742	3.794	0.805
Native college	11.914	4.078	0.096	2.252	16.981
Immigrant noncollege	3.358	1.442	0.603	1.884	4.064
Immigrant college	0.397	1.824	1.885	3.787	8.150

Table A.30: Probability of marriage, male born in '70s living in West

Male Human Capital	Female Human Capital Native noncollege	 Native college	 Immigrant noncollege	 Immigrant college	 Single
Simulation					
Native noncollege	0.598	0.055	0.060	0.008	0.280
Native college	0.375	0.322	0.031	0.036	0.235
Immigrant noncollege	0.203	0.027	0.487	0.043	0.241
Immigrant college	0.102	0.147	0.165	0.311	0.276
Data					
Native noncollege	0.627	0.062	0.067	0.010	0.233
Native college	0.358	0.362	0.034	0.043	0.203
Immigrant noncollege	0.217	0.016	0.620	0.033	0.114
Immigrant college	0.154	0.131	0.252	0.319	0.144
Diff in SE					
Native noncollege	4.213	2.102	1.968	1.631	7.564
Native college	1.501	3.386	0.796	1.380	3.315
Immigrant noncollege	1.387	3.220	10.802	2.194	15.785
Immigrant college	2.308	0.749	3.181	0.277	5.952

Table A.31: Probability of marriage, female born in '50s living in North

Female Human Capital	Male Human Capital Native noncollege	Native college	Immigrant noncollege	Immigrant college	Single
Simulation					
Native noncollege	0.718	0.150	0.035	0.005	0.092
Native college	0.238	0.483	0.018	0.026	0.234
Immigrant noncollege	0.314	0.034	0.562	0.043	0.048
Immigrant college	0.149	0.182	0.179	0.364	0.127
Data					
Native noncollege	0.735	0.131	0.030	0.005	0.099
Native college	0.201	0.539	0.012	0.032	0.216
Immigrant noncollege	0.355	0.058	0.374	0.123	0.090
Immigrant college	0.135	0.240	0.036	0.435	0.153
Diff in SE					
Native noncollege	2.339	3.452	1.951	0.670	1.397
Native college	2.753	3.293	1.800	0.936	1.240
Immigrant noncollege	1.772	2.152	7.962	4.985	3.026
Immigrant college	0.407	1.383	7.738	1.457	0.750

Table A.32: Probability of marriage, female born in '50s living in South

Female Human Capital	Male Human Capital: Native noncollege	Native college	Immigrant noncollege	Immigrant college	Single
Simulation					
Native noncollege	0.714	0.170	0.043	0.007	0.066
Native college	0.224	0.568	0.021	0.038	0.148
Immigrant noncollege	0.264	0.040	0.608	0.061	0.028
Immigrant college	0.105	0.189	0.162	0.510	0.034
Data					
Native noncollege	0.729	0.154	0.041	0.007	0.069
Native college	0.194	0.595	0.021	0.037	0.153
Immigrant noncollege	0.276	0.070	0.492	0.100	0.063
Immigrant college	0.115	0.272	0.061	0.462	0.089
Diff in SE					
Native noncollege	3.028	3.949	0.835	0.024	0.974
Native college	2.833	1.964	0.022	0.187	0.537
Immigrant noncollege	1.005	4.349	8.563	4.818	5.278
Immigrant college	0.474	2.862	6.423	1.458	2.946

Table A.33: Probability of marriage, female born in '50s living in West

Female Human Capital	Male Human Capital: Native noncollege	Native college	Immigrant noncollege	Immigrant college	Single
Simulation					
Native noncollege	0.707	0.163	0.037	0.005	0.088
Native college	0.226	0.539	0.018	0.029	0.188
Immigrant noncollege	0.293	0.039	0.573	0.047	0.048
Immigrant college	0.136	0.207	0.175	0.395	0.086
Data					
Native noncollege	0.751	0.138	0.037	0.005	0.069
Native college	0.202	0.587	0.019	0.032	0.161
Immigrant noncollege	0.289	0.051	0.489	0.101	0.069
Immigrant college	0.142	0.223	0.048	0.486	0.100
Diff in SE					
Native noncollege	7.987	5.707	0.230	0.394	5.796
Native college	1.762	2.965	0.113	0.449	2.303
Immigrant noncollege	0.201	1.624	5.004	5.382	2.404
Immigrant college	0.204	0.459	6.815	2.084	0.525

Table A.34: Probability of marriage, female born in '60s living in North

Female Human Capital	Male Human Capital: Native noncollege	Native college	Immigrant noncollege	Immigrant college	Single
Simulation					
Native noncollege	0.665	0.158	0.035	0.005	0.138
Native college	0.234	0.444	0.021	0.024	0.279
Immigrant noncollege	0.292	0.043	0.554	0.045	0.066
Immigrant college	0.159	0.171	0.192	0.320	0.158
Data					
Native noncollege	0.677	0.125	0.030	0.005	0.163
Native college	0.213	0.491	0.016	0.021	0.259
Immigrant noncollege	0.281	0.064	0.442	0.102	0.111
Immigrant college	0.171	0.275	0.048	0.331	0.175
Diff in SE					
Native noncollege	1.872	7.339	2.168	0.286	5.144
Native college	1.774	3.338	1.169	0.662	1.596
Immigrant noncollege	0.669	2.249	5.766	4.854	3.642
Immigrant college	0.421	3.036	8.741	0.303	0.566

Table A.35: Probability of marriage, female born in '60s living in South

Female Human Capital	Male Human Capital: Native noncollege	Native college	Immigrant noncollege	Immigrant college	Single
Simulation					
Native noncollege	0.667	0.183	0.043	0.007	0.099
Native college	0.223	0.539	0.024	0.034	0.179
Immigrant noncollege	0.246	0.051	0.602	0.062	0.039
Immigrant college	0.118	0.190	0.182	0.465	0.045
Data					
Native noncollege	0.679	0.165	0.038	0.009	0.109
Native college	0.200	0.561	0.018	0.030	0.191
Immigrant noncollege	0.254	0.070	0.518	0.094	0.064
Immigrant college	0.147	0.314	0.067	0.372	0.100
Diff in SE					
Native noncollege	2.691	5.199	2.648	1.319	3.393
Native college	2.675	2.023	1.919	1.309	1.392
Immigrant noncollege	0.760	3.276	7.384	4.854	4.525
Immigrant college	1.630	5.282	9.115	3.800	3.612

Table A.36: Probability of marriage, female born in '60s living in West

Female Human Capital	Male Human Capital: Native noncollege	Native college	Immigrant noncollege	Immigrant college	Single
Simulation					
Native noncollege	0.654	0.173	0.036	0.005	0.131
Native college	0.222	0.503	0.020	0.026	0.228
Immigrant noncollege	0.274	0.050	0.561	0.049	0.067
Immigrant college	0.147	0.199	0.189	0.355	0.111
Data					
Native noncollege	0.706	0.145	0.035	0.006	0.108
Native college	0.233	0.538	0.018	0.032	0.179
Immigrant noncollege	0.267	0.050	0.516	0.090	0.077
Immigrant college	0.168	0.276	0.064	0.398	0.094
Diff in SE					
Native noncollege	10.261	7.104	0.621	0.877	6.868
Native college	0.955	2.698	0.687	1.153	4.856
Immigrant noncollege	0.568	0.013	3.189	5.159	1.370
Immigrant college	0.760	2.287	6.728	1.165	0.771

Table A.37: Probability of marriage, female born in '70s living in North

Female Human Capital	Male Human Capital: Native noncollege	Native college	Immigrant noncollege	Immigrant college	Single
Simulation					
Native noncollege	0.635	0.147	0.055	0.006	0.157
Native college	0.199	0.404	0.025	0.030	0.342
Immigrant noncollege	0.272	0.046	0.554	0.042	0.087
Immigrant college	0.125	0.193	0.179	0.273	0.230
Data					
Native noncollege	0.613	0.121	0.047	0.006	0.213
Native college	0.208	0.451	0.019	0.029	0.292
Immigrant noncollege	0.253	0.044	0.503	0.075	0.125
Immigrant college	0.142	0.269	0.062	0.315	0.212
Diff in SE					
Native noncollege	2.630	4.887	2.348	0.105	8.204
Native college	0.744	3.241	1.561	0.010	3.740
Immigrant noncollege	1.248	0.246	2.914	3.606	3.332
Immigrant college	0.753	2.610	7.362	1.351	0.671

Table A.38: Probability of marriage, female born in '70s living in South

Female Human Capital	Male Human Capital: Native noncollege	Native college	Immigrant noncollege	Immigrant college	Single
Simulation					
Native noncollege	0.635	0.175	0.067	0.010	0.113
Native college	0.193	0.509	0.030	0.045	0.224
Immigrant noncollege	0.231	0.057	0.601	0.060	0.051
Immigrant college	0.095	0.231	0.175	0.429	0.070
Data					
Native noncollege	0.618	0.152	0.070	0.009	0.151
Native college	0.208	0.509	0.022	0.038	0.224
Immigrant noncollege	0.277	0.071	0.483	0.084	0.085
Immigrant college	0.151	0.298	0.060	0.345	0.146
Diff in SE					
Native noncollege	2.843	5.450	0.808	0.136	8.776
Native college	1.695	0.020	2.495	1.754	0.042
Immigrant noncollege	4.712	2.525	10.826	3.956	5.575
Immigrant college	3.646	3.425	11.310	4.147	5.052

Table A.39: Probability of marriage, female born in '70s living in West

Female Human Capital	Male Human Capital: Native noncollege	Native college	Immigrant noncollege	Immigrant college	Single
Simulation					
Native noncollege	0.624	0.162	0.057	0.007	0.151
Native college	0.191	0.464	0.025	0.034	0.285
Immigrant noncollege	0.255	0.054	0.557	0.046	0.088
Immigrant college	0.116	0.230	0.176	0.313	0.166
Data					
Native noncollege	0.647	0.128	0.071	0.009	0.144
Native college	0.242	0.490	0.020	0.030	0.217
Immigrant noncollege	0.235	0.042	0.583	0.068	0.072
Immigrant college	0.177	0.268	0.078	0.337	0.140
Diff in SE					
Native noncollege	3.404	6.867	3.819	1.744	1.368
Native college	4.218	1.844	1.203	0.725	5.907
Immigrant noncollege	1.858	2.425	2.132	3.446	2.498
Immigrant college	2.520	1.369	5.776	0.821	1.185

Table A.40: Probability of region choice, male born in '50s

Human Capital	Simulation	Data	Data SE	Diff in SE
North				
Native noncollege	0.244	0.241	0.003	1.197
Native college	0.227	0.244	0.006	2.793
Immigrant noncollege	0.205	0.177	0.007	3.845
Immigrant college	0.211	0.227	0.017	0.902
South				
Native noncollege	0.407	0.410	0.004	0.877
Native college	0.415	0.435	0.007	3.040
Immigrant noncollege	0.456	0.480	0.009	2.560
Immigrant college	0.473	0.468	0.021	0.239
West				
Native noncollege	0.349	0.349	0.004	0.168
Native college	0.358	0.321	0.006	5.796
Immigrant noncollege	0.339	0.343	0.009	0.397
Immigrant college	0.316	0.305	0.019	0.562

Table A.41: Probability of region choice, male born in '60s

Human Capital	Simulation	Data	Data SE	Diff in SE
North				
Native noncollege	0.246	0.251	0.003	1.894
Native college	0.229	0.233	0.005	0.702
Immigrant noncollege	0.206	0.195	0.006	1.754
Immigrant college	0.212	0.221	0.014	0.625
South				
Native noncollege	0.411	0.412	0.003	0.323
Native college	0.417	0.461	0.006	7.932
Immigrant noncollege	0.465	0.465	0.008	0.023
Immigrant college	0.476	0.497	0.017	1.236
West				
Native noncollege	0.343	0.336	0.003	2.076
Native college	0.353	0.306	0.005	9.226
Immigrant noncollege	0.329	0.340	0.007	1.492
Immigrant college	0.312	0.282	0.016	1.950

Table A.42: Probability of region choice, male born in '70s

Human Capital	Simulation	Data	Data SE	Diff in SE
North				
Native noncollege	0.249	0.262	0.004	3.656
Native college	0.233	0.245	0.006	2.260
Immigrant noncollege	0.210	0.209	0.006	0.126
Immigrant college	0.215	0.240	0.014	1.793
South				
Native noncollege	0.412	0.407	0.004	1.161
Native college	0.415	0.466	0.006	7.934
Immigrant noncollege	0.467	0.419	0.007	6.493
Immigrant college	0.473	0.486	0.016	0.799
West				
Native noncollege	0.339	0.331	0.004	2.203
Native college	0.352	0.289	0.006	10.878
Immigrant noncollege	0.323	0.372	0.007	6.733
Immigrant college	0.312	0.274	0.015	2.611

Table A.43: Probability of region choice, female born in '50s

Human Capital	Simulation	Data	Data SE	Diff in SE
North				
Native noncollege	0.242	0.229	0.003	4.100
Native college	0.228	0.285	0.008	7.020
Immigrant noncollege	0.199	0.176	0.007	3.184
Immigrant college	0.212	0.235	0.020	1.192
South				
Native noncollege	0.415	0.420	0.004	1.439
Native college	0.422	0.402	0.009	2.219
Immigrant noncollege	0.469	0.487	0.010	1.888
Immigrant college	0.480	0.464	0.023	0.682
West				
Native noncollege	0.343	0.350	0.004	2.123
Native college	0.350	0.313	0.008	4.487
Immigrant noncollege	0.332	0.337	0.009	0.568
Immigrant college	0.308	0.300	0.021	0.361

Table A.44: Probability of region choice, female born in '60s

Human Capital	Simulation	Data	Data SE	Diff in SE
North				
Native noncollege	0.243	0.237	0.003	2.275
Native college	0.230	0.274	0.006	6.801
Immigrant noncollege	0.201	0.183	0.006	2.966
Immigrant college	0.215	0.248	0.016	2.105
South				
Native noncollege	0.421	0.427	0.003	1.902
Native college	0.426	0.424	0.007	0.265
Immigrant noncollege	0.479	0.486	0.008	0.784
Immigrant college	0.480	0.499	0.018	1.042
West				
Native noncollege	0.336	0.336	0.003	0.057
Native college	0.344	0.302	0.007	6.319
Immigrant noncollege	0.320	0.332	0.008	1.602
Immigrant college	0.305	0.252	0.016	3.293

Table A.45: Probability of region choice, female born in '70s

Human Capital	Simulation	Data	Data SE	Diff in SE
North				
Native noncollege	0.246	0.253	0.004	2.005
Native college	0.234	0.269	0.007	5.381
Immigrant noncollege	0.205	0.199	0.006	0.884
Immigrant college	0.218	0.248	0.014	2.222
South				
Native noncollege	0.423	0.420	0.004	0.588
Native college	0.427	0.434	0.007	0.983
Immigrant noncollege	0.481	0.429	0.007	7.031
Immigrant college	0.478	0.498	0.016	1.288
West				
Native noncollege	0.331	0.327	0.004	1.239
Native college	0.340	0.297	0.007	6.292
Immigrant noncollege	0.314	0.371	0.007	7.932
Immigrant college	0.304	0.254	0.014	3.685

B. APPENDIX TO CHAPTER 2

Data and descriptives

Table B.1: Description of parenting dimensions in the LSAC

Dimension	Description
Parental warmth	Parent shows affection with hugs, kisses and holds the child often, hugs the child without a reason, expresses happiness about child, has warm and close times with the child, enjoys listening to child and doing things with them, parent feels close to child when it is happy or upset
Parental hostility	Frequency with which parents react to child's behaviour with praise or disapproval, parents react with anger when punishing child, feel to have problems managing child
Parental consistency	Frequency of making sure child completes requests, punishment if child does not complete requests, how often child gets away with things which parents feel they should be punished for, child gets out of punishment or ignores it
Parental reasoning	Frequency with which parent explains why child gets corrected, reasons about misbehaviour and why rules should be obeyed, explains consequences of behaviour, emphasizes reasons for rules

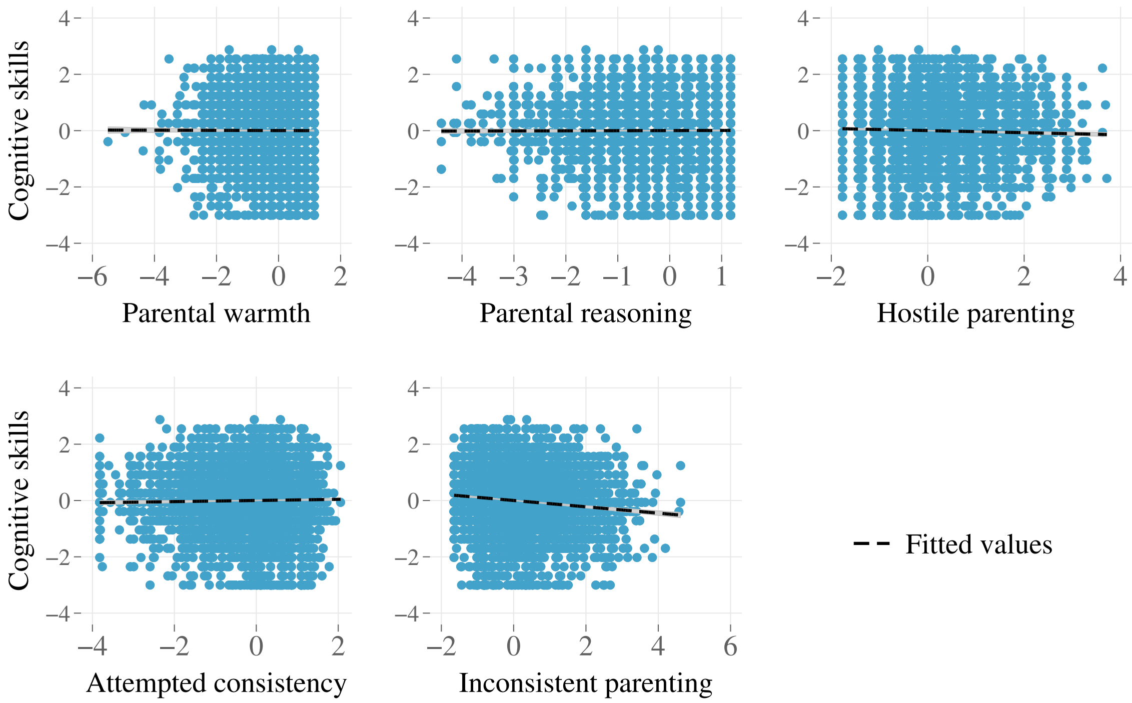


Figure B.1: Correlation of parenting dimensions with cognitive skills

Notes: The figure displays the relationship between cognitive skills (measured by the MRT) and different parenting styles. Each data point represents a child from the 8-9 age group. In addition to the data points, a line is plotted on the graph, which represents the fitted values based on a linear regression analysis. The line slope is estimated using population weights.

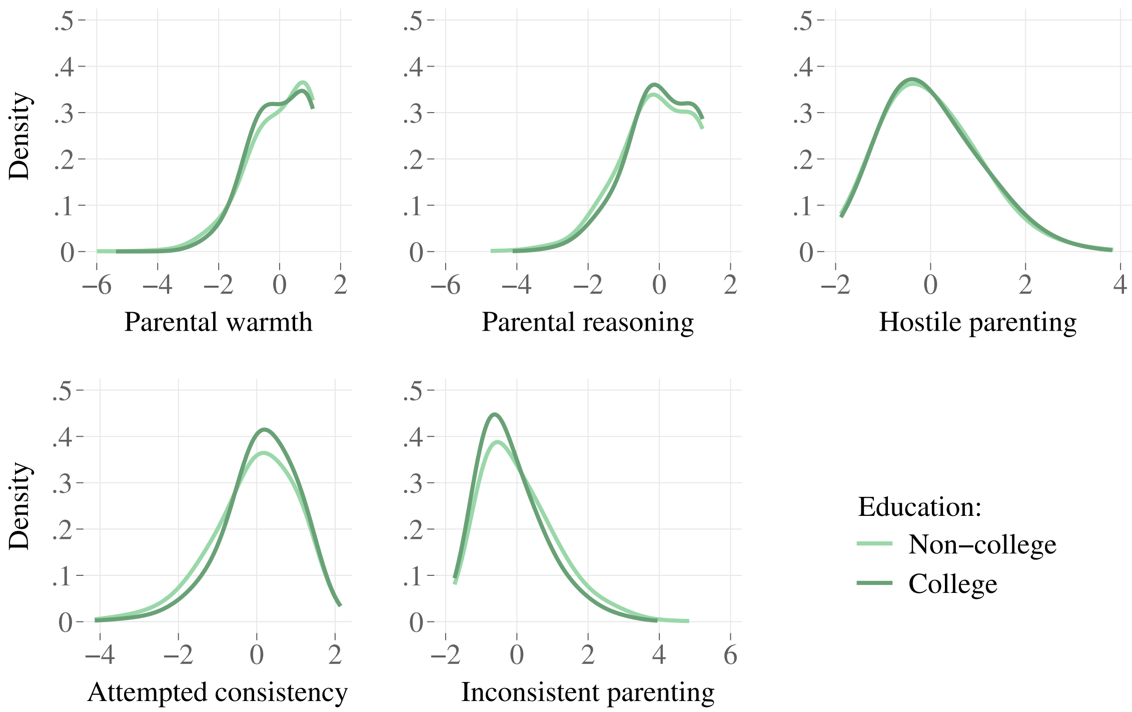


Figure B.2: Distribution of parenting dimensions by primary care giver's education

Notes: The figure displays the empirical distribution (smoothed using the kernel function approach with population weights) of different parenting styles by primary care giver education for children aged 8-9.

Table B.2: Rotated factor loadings for single factors

	Age:					
	4-5	6-7	8-9	10-11	12-13	14-15
Parental warmth:						
Expresses affection	**0.739**	**0.817**	**0.829**	**0.835**	**0.850**	**0.848**
Hugs child	**0.741**	**0.775**	**0.776**	**0.792**	**0.805**	**0.795**
Expresses happiness	**0.757**	**0.771**	**0.796**	**0.796**	**0.790**	**0.794**
Warm/close times together	**0.797**	**0.829**	**0.850**	**0.850**	**0.847**	**0.843**
Enjoy time together	**0.747**	**0.786**	**0.812**	**0.795**	**0.792**	**0.801**
Feels close to child	**0.753**	**0.796**	**0.796**	**0.803**	**0.800**	**0.793**
Parental hostility:						
Praise child	**-0.550**	**-0.555**	**-0.641**	**-0.649**	**-0.688**	**-0.711**
Disapproval	**0.731**	**0.754**	**0.763**	**0.780**	**0.805**	**0.804**
Angry when punishing	**0.673**	**0.678**	**0.659**	**0.692**	**0.676**	**0.682**
Having problems managing	**0.743**	**0.744**	**0.733**	**0.752**	**0.760**	**0.756**
Parental consistency: Factor 1						
Ensures requests complete	-0.053	-0.055	-0.035	-0.043	-0.031	-0.050
Punishes child	-0.245	-0.223	**-0.279**	**-0.263**	-0.232	-0.188
Child gets away	**0.779**	**0.771**	**0.774**	**0.802**	**0.805**	**0.828**
Child gets out of punishment	**0.804**	**0.800**	**0.815**	**0.809**	**0.816**	**0.824**
Child ignores punishment	**0.793**	**0.812**	**0.800**	**0.808**	**0.818**	**0.842**
Parental consistency: Factor 2						
Ensures requests complete		**0.847**	**0.860**	**0.864**	**0.853**	**0.838**
Punishes child		**0.779**	**0.750**	**0.771**	**0.778**	**0.787**
Child gets away		**-0.259**	**-0.259**	-0.204	-0.202	-0.166
Child gets out of punishment		-0.147	-0.124	-0.144	-0.123	-0.131
Child ignores punishment		-0.021	-0.038	-0.060	-0.035	-0.039
Parental inductive reasoning:						
Explains corrections	**0.870**	**0.887**	**0.881**	**0.887**	**0.897**	**0.904**
Reasons when misbehaves	**0.870**	**0.819**	**0.751**	**0.738**	**0.756**	**0.746**
Reasons for rules		**0.882**	**0.867**	**0.864**	**0.882**	**0.887**
Explains consequences			**0.892**	**0.896**	**0.913**	**0.906**
Emphasizes reasons			**0.888**	**0.894**	**0.905**	**0.907**

Notes: Factor loadings larger than 0.25 in absolute value printed in bold. To summarize the variation of all measures, one factor was sufficient expect for parental consistency from wave 4 onward. Eigenvalues of bigger than 1 indicated which factors to include in the analysis.

Table B.3: Rotated factor loadings for joint analysis

	Age:					
	4-5	6-7	8-9	10-11	12-13	14-15
Factor 1:						
Parental warmth	**0.849**	**0.730**	**0.693**	**-0.544**	**-0.645**	**-0.668**
Hostile parenting	**-0.260**	-0.152	-0.103	**0.876**	**0.883**	**0.886**
Attempted consistency	0.055	**0.637**	**0.637**	0.193	0.108	0.075
Parental inconsistency		0.030	0.037	**0.766**	**0.767**	**0.787**
Parental reasoning	**0.856**	**0.848**	**0.856**	-0.041	0.011	0.053
Factor 2:						
Parental warmth		**-0.352**	**-0.433**	**0.612**	**0.513**	**0.480**
Hostile parenting		**0.851**	**0.859**	-0.049	0.073	0.116
Attempted consistency		0.201	0.213	**0.664**	**0.703**	**0.697**
Parental inconsistency		**0.808**	**0.778**	0.059	0.084	0.094
Parental reasonig		-0.039	-0.017	**0.854**	**0.859**	**0.860**

Notes: Factor loadings larger than 0.25 in absolute value printed in bold. To summarize the variation of all measures, two factors were sufficient to summarize the data expect for wave 3. Eigenvalues of bigger than 1 indicated which factors to include in the analysis. From wave 6 factor 1 is factor 2 and wise versa which is why we swap them in the data to get consistent measures across waves.

Table B.4: Rotated factor loadings at age 4-5 - joint estimation

	Factor 1	Factor 2	Factor 3	Factor 4
Parental warmth:				
Expresses affection	**0.808**	-0.070	0.024	0.014
Hugs child	**0.812**	-0.061	0.016	0.024
Expresses happiness	**0.660**	0.001	-0.157	**0.335**
Warm/close times together	**0.742**	0.012	-0.118	0.238
Enjoy time together	**0.609**	0.001	-0.225	**0.357**
Feels close to child	**0.649**	-0.003	-0.217	**0.277**
Parental hostility:				
Praise child	**0.276**	0.076	**-0.468**	**0.319**
Disapproval	-0.143	0.122	**0.686**	0.038
Angry when punishing	-0.020	0.077	**0.684**	-0.050
Having problems managing	-0.113	**0.289**	**0.674**	-0.034
Parental consistency:				
Ensures requests complete	0.053	**-0.424**	0.062	**0.479**
Punishes child	0.023	**-0.668**	**0.286**	**0.257**
Child gets away	-0.018	**0.747**	0.212	-0.021
Child gets out of punishment	-0.020	**0.763**	0.176	0.037
Child ignores punishment	-0.077	**0.621**	**0.437**	0.019
Parental inductive reasoning:				
Explains corrections	0.212	-0.072	-0.011	**0.761**
Reasons when misbehaves	**0.256**	-0.016	-0.033	**0.741**

Notes: Factor loadings larger than 0.25 in absolute value printed in bold. Factors can be assigned the following across waves: factor 1: parental warmth, factor 2: reasoning, factor 3: hostile and inconsistent parenting, factor 4: consistency. As in wave 3 factor 2 describes inconsistency, when assign it to the variable consistency, but we reverse values of factor 2 before assignment to ensure comparability across waves. Instead we assign factor 4 as reasoning.

B.1. Estimation tables

B.1.1. Non-cognitive skills

Table B.5: Rotated factor loadings at age 6-7 - joint estimation

	Factor 1	Factor 2	Factor 3	Factor 4
Parental warmth:				
Expresses affection	**0.837**	0.104	0.002	-0.057
Hugs child	**0.799**	0.121	0.021	-0.057
Expresses happiness	**0.677**	**0.336**	-0.188	0.011
Warm/close times together	**0.787**	0.234	-0.119	-0.014
Enjoy time together	**0.701**	**0.268**	-0.172	-0.031
Feels close to child	**0.736**	0.213	-0.170	-0.036
Parental hostility:				
Praise child	**0.400**	0.162	**-0.449**	0.110
Disapproval	-0.197	0.038	**0.699**	0.065
Angry when punishing	-0.039	-0.083	**0.687**	0.011
Having problems managing	-0.150	0.021	**0.686**	**0.272**
Parental consistency:				
Ensures requests complete	0.151	**0.254**	0.105	**-0.537**
Punishes child	0.037	0.198	0.218	**-0.722**
Child gets away	-0.024	-0.044	**0.330**	**0.715**
Child gets out of punishment	0.003	0.003	**0.298**	**0.691**
Child ignores punishment	-0.059	0.012	**0.534**	**0.543**
Parental inductive reasoning:				
Explains corrections	0.196	**0.866**	-0.023	-0.077
Reasons when misbehaves	**0.285**	**0.736**	0.007	-0.067
Reasons for rules	0.212	**0.855**	-0.003	-0.068

Notes: Factor loadings larger than 0.25 in absolute value printed in bold. Factors can be assigned the following across waves: factor 1: parental warmth, factor 2: reasoning, factor 3: hostile and inconsistent parenting, factor 4: consistency. As in wave 4 factor 4 describes inconsistency, when creating the variable consistency, we reverse values of factor 4 before assignment to ensure comparability across waves.

Table B.6: Rotated factor loadings at age 8-9 - joint estimation

	Factor 1	Factor 2	Factor 3	Factor 4
Parental warmth:				
Expresses affection	**0.833**	0.155	0.001	-0.053
Hugs child	**0.794**	0.144	0.009	-0.055
Expresses happiness	**0.705**	**0.319**	-0.179	0.034
Warm/close times together	**0.787**	**0.269**	-0.114	-0.014
Enjoy time together	**0.729**	**0.256**	-0.179	-0.048
Feels close to child	**0.736**	0.212	-0.191	-0.043
Parental hostility:				
Praise child	**0.437**	0.121	**-0.479**	0.067
Disapproval	**-0.268**	0.088	**0.680**	0.020
Angry when punishing	-0.047	-0.025	**0.690**	0.008
Having problems managing	-0.163	0.059	**0.688**	**0.254**
Parental consistency:				
Ensures requests complete	0.139	0.219	0.132	**-0.549**
Punishes child	0.020	0.181	0.203	**-0.747**
Child gets away	-0.027	-0.050	**0.355**	**0.698**
Child gets out of punishment	-0.000	-0.028	**0.339**	**0.672**
Child ignores punishment	-0.072	0.012	**0.536**	**0.538**
Parental inductive reasoning:				
Explains corrections	0.170	**0.864**	-0.017	-0.073
Reasons when misbehaves	**0.271**	**0.687**	0.033	-0.107
Reasons for rules	0.180	**0.852**	0.002	-0.047
Explains consequences	0.202	**0.864**	0.037	-0.061
Emphasizes reasons	0.171	**0.874**	0.003	-0.035

Notes: Factor loadings larger than 0.25 in absolute value printed in bold. Factors can be assigned the following across waves: factor 1: parental warmth, factor 2: reasoning, factor 3: hostile and inconsistent parenting, factor 4: consistency. As in wave 5 factor 4 describes inconsistency, when creating the variable consistency, we reverse values of factor 4 before assignment to ensure comparability across waves.

Table B.7: Rotated factor loadings at age 10-11 - joint estimation

	Factor 1	Factor 2	Factor 3	Factor 4
Parental warmth:				
Expresses affection	**0.840**	0.156	-0.009	-0.066
Hugs child	**0.817**	0.121	0.002	-0.076
Expresses happiness	**0.724**	**0.277**	-0.185	0.021
Warm/close times together	**0.792**	0.245	-0.147	-0.019
Enjoy time together	**0.697**	**0.261**	-0.237	-0.022
Feels close to child	**0.724**	0.196	-0.235	-0.005
Parental hostility:				
Praise child	**0.454**	0.085	**-0.474**	0.051
Disapproval	**-0.291**	0.117	**0.690**	-0.041
Angry when punishing	-0.103	0.002	**0.689**	-0.005
Having problems managing	-0.191	0.046	**0.716**	0.167
Parental consistency:				
Ensures requests complete	0.114	0.208	0.109	**-0.598**
Punishes child	0.031	0.187	0.124	**-0.772**
Child gets away	-0.024	-0.015	**0.445**	**0.653**
Child gets out of punishment	-0.000	-0.018	**0.415**	**0.630**
Child ignores punishment	-0.082	-0.012	**0.581**	**0.488**
Parental inductive reasoning:				
Explains corrections	0.151	**0.876**	-0.007	-0.076
Reasons when misbehaves	**0.258**	**0.693**	0.013	-0.088
Reasons for rules	0.149	**0.861**	-0.016	-0.026
Explains consequences	0.185	**0.873**	0.057	-0.079
Emphasizes reasons	0.161	**0.883**	0.020	-0.050

Notes: Factor loadings larger than 0.25 in absolute value printed in bold. Factors can be assigned the following across waves: factor 1: parental warmth, factor 2: reasoning, factor 3: hostile and inconsistent parenting, factor 4: consistency. As in wave 6 factor 4 describes inconsistency, when creating the variable consistency, we reverse values of factor 4 before assignment to ensure comparability across waves.

Table B.8: Rotated factor loadings at age 12-13 - joint estimation

	Factor 1	Factor 2	Factor 3	Factor 4
Parental warmth:				
Expresses affection	**0.849**	0.140	-0.021	0.086
Hugs child	**0.820**	0.113	-0.002	0.106
Expresses happiness	**0.715**	**0.260**	-0.203	-0.084
Warm/close times together	**0.794**	0.210	-0.148	0.013
Enjoy time together	**0.715**	0.192	-0.232	0.024
Feels close to child	**0.728**	0.163	**-0.258**	0.007
Parental hostility:				
Praise child	**0.482**	0.037	**-0.467**	-0.154
Disapproval	**-0.311**	0.166	**0.661**	0.159
Angry when punishing	-0.099	0.034	**0.684**	0.156
Having problems managing	-0.236	0.113	**0.716**	-0.073
Parental consistency:				
Ensures requests complete	0.108	0.249	0.030	**0.595**
Punishes child	0.026	0.196	0.004	**0.766**
Child gets away	-0.091	0.009	**0.584**	**-0.531**
Child gets out of punishment	0.005	0.002	**0.549**	**-0.531**
Child ignores punishment	-0.130	0.042	**0.675**	**-0.368**
Parental inductive reasoning:				
Explains corrections	0.121	**0.889**	0.012	0.070
Reasons when misbehaves	0.240	**0.707**	0.021	0.076
Reasons for rules	0.129	**0.877**	0.034	0.017
Explains consequences	0.154	**0.890**	0.065	0.092
Emphasizes reasons	0.126	**0.898**	0.047	0.048

Notes: Factor loadings larger than 0.25 in absolute value printed in bold. Factors can be assigned the following across waves: factor 1: parental warmth, factor 2: reasoning, factor 3: hostile and inconsistent parenting, factor 4: consistency.

Table B.9: Rotated factor loadings at age 14-15 - joint estimation

	Factor 1	Factor 2	Factor 3	Factor 4
Parental warmth:				
Expresses affection	**0.854**	0.105	-0.023	0.110
Hugs child	**0.818**	0.083	-0.009	0.121
Expresses happiness	**0.745**	0.235	-0.153	-0.074
Warm/close times together	**0.792**	0.193	-0.155	-0.032
Enjoy time together	**0.721**	0.180	-0.222	-0.039
Feels close to child	**0.725**	0.152	-0.246	-0.067
Parental hostility:				
Praise child	**0.522**	-0.006	**-0.407**	-0.201
Disapproval	**-0.347**	0.226	**0.568**	**0.294**
Angry when punishing	-0.121	0.083	**0.600**	**0.322**
Having problems managing	-0.239	0.121	**0.718**	0.097
Parental consistency:				
Ensures requests complete	0.082	0.223	-0.134	**0.640**
Punishes child	0.011	0.206	-0.197	**0.736**
Child gets away	-0.094	0.037	**0.744**	**-0.292**
Child gets out of punishment	-0.042	0.020	**0.706**	**-0.313**
Child ignores punishment	-0.148	0.052	**0.771**	-0.166
Parental inductive reasoning:				
Explains corrections	0.100	**0.904**	0.022	0.055
Reasons when misbehaves	0.230	**0.713**	0.052	0.122
Reasons for rules	0.102	**0.890**	0.033	0.021
Explains consequences	0.138	**0.889**	0.072	0.091
Emphasizes reasons	0.101	**0.899**	0.066	0.076

Notes: Factor loadings larger than 0.25 in absolute value printed in bold. Factors can be assigned the following across waves: factor 1: parental warmth, factor 2: reasoning, factor 3: hostile and inconsistent parenting, factor 4: consistency.

Table B.10: Estimated parameters of production function for non-cognitive skills at age 10-11

	OLS	VA	FE	AB	CU	CV
Parental warmth	0.050***	0.011	0.036***	0.032**	0.037	0.038
	(0.015)	(0.012)	(0.013)	(0.013)	(0.030)	(0.025)
Parental reasoning	-0.084***	-0.026***	-0.033***	-0.035***	-0.059***	-0.043**
	(0.012)	(0.009)	(0.010)	(0.010)	(0.023)	(0.018)
Hostile parenting	-0.382***	-0.177***	-0.170***	-0.161***	-0.303***	-0.213***
	(0.014)	(0.012)	(0.013)	(0.013)	(0.028)	(0.023)
Inconsistent parenting	-0.145***	-0.068***	-0.053***	-0.077***	-0.129***	-0.110***
	(0.014)	(0.011)	(0.013)	(0.013)	(0.029)	(0.022)
Attempted consistency	0.015	-0.002	0.006	0.006	-0.010	-0.009
	(0.011)	(0.009)	(0.010)	(0.010)	(0.021)	(0.017)
Educational time parents	-0.001	-0.002	-0.000	-0.001	-0.002	0.001
	(0.003)	(0.002)	(0.002)	(0.002)	(0.005)	(0.003)
Educational time others	-0.001	-0.000	-0.003	0.003	0.015*	0.011
	(0.005)	(0.004)	(0.005)	(0.004)	(0.009)	(0.007)
Care time parents	-0.002	-0.001	-0.000	-0.001	0.000	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
Care time others	0.007***	0.004**	0.001	0.002	-0.001	0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)
Lagged test outcome		0.634***		0.253***		0.646***
		(0.013)		(0.019)		(0.021)
Observations	7,299	6,703	6,599	6,463	2,267	2,264
Lagged dependent	NO	YES	NO	YES	NO	YES
Lagged inputs	NO	NO	NO	NO	YES	YES
Fixed effects	NO	NO	YES	YES	NO	NO

Notes: Each specification includes controls as the age of the primary caregiver, the number of siblings, the log of family income, and dummies for the college education of the primary caregiver, the presence of both biological parents at home, the gender of the study child, the indigenous status of the study child, the study child's cohort, and the day of data collection.

Table B.11: Estimated parameters of production function for non-cognitive skills at age 12-13

	OLS	VA	FE	AB	CU	CV
Parental warmth	0.052***	0.028**	0.028**	0.044***	0.077**	0.064**
	(0.016)	(0.012)	(0.012)	(0.013)	(0.032)	(0.025)
Parental reasoning	-0.082***	-0.035***	-0.030***	-0.026***	-0.052**	-0.025
	(0.012)	(0.009)	(0.009)	(0.010)	(0.024)	(0.018)
Hostile parenting	-0.386***	-0.184***	-0.178***	-0.180***	-0.316***	-0.267***
	(0.017)	(0.013)	(0.013)	(0.014)	(0.030)	(0.026)
Inconsistent parenting	-0.121***	-0.053***	-0.065***	-0.059***	-0.055*	-0.054**
	(0.016)	(0.012)	(0.012)	(0.012)	(0.031)	(0.024)
Attempted consistency	0.025*	0.009	0.002	0.005	0.008	0.010
	(0.013)	(0.009)	(0.010)	(0.010)	(0.021)	(0.017)
Educational time parents	-0.004*	-0.000	0.000	0.001	-0.008*	-0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.004)	(0.003)
Educational time others	-0.006	-0.003	-0.004	-0.005	-0.012	-0.011
	(0.005)	(0.004)	(0.004)	(0.004)	(0.012)	(0.009)
Care time parents	0.001	0.000	0.001	-0.000	0.000	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
Care time others	0.000	-0.001	0.001	0.000	0.003	0.000
	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
Lagged test outcome		0.658***		0.253***		0.651***
		(0.013)		(0.019)		(0.022)
Observations	6,544	6,346	6,599	6,463	2,067	2,066
Lagged dependent	NO	YES	NO	YES	NO	YES
Lagged inputs	NO	NO	NO	NO	YES	YES
Fixed effects	NO	NO	YES	YES	NO	NO

Notes: Each specification includes controls as the age of the primary caregiver, the number of siblings, the log of family income, and dummies for the college education of the primary caregiver, the presence of both biological parents at home, the gender of the study child, the indigenous status of the study child, the study child's cohort, and the day of data collection.

Table B.12: Estimated parameters of production function for non-cognitive skills at age 14-15

	OLS	VA	FE	AB	CU	CV
Parental warmth	0.054***	0.019	0.032**	0.036**	-0.020	0.002
	(0.017)	(0.013)	(0.014)	(0.014)	(0.039)	(0.028)
Parental reasoning	-0.086***	-0.037***	-0.040***	-0.033***	-0.041*	-0.044**
	(0.013)	(0.010)	(0.011)	(0.011)	(0.024)	(0.019)
Hostile parenting	-0.363***	-0.179***	-0.167***	-0.158***	-0.253***	-0.216***
	(0.019)	(0.015)	(0.016)	(0.015)	(0.036)	(0.028)
Inconsistent parenting	-0.157***	-0.085***	-0.103***	-0.080***	-0.175***	-0.135***
	(0.018)	(0.014)	(0.015)	(0.015)	(0.038)	(0.031)
Attempted consistency	0.024*	0.004	0.025**	0.016	0.017	-0.005
	(0.013)	(0.011)	(0.011)	(0.011)	(0.024)	(0.020)
Educational time parents	0.001	0.002	0.002	0.002	0.005*	0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)
Educational time others	0.002	0.000	-0.002	-0.002	0.003	0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.004)	(0.003)
Care time parents	-0.001	-0.001	-0.001	-0.001	0.001	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
Care time others	-0.001	-0.001	-0.001	-0.001	-0.002	-0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)
Lagged test outcome		0.635***		0.253***		0.653***
		(0.014)		(0.019)		(0.023)
Observations	5,726	5,531	6,599	6,463	1,753	1,753
Lagged dependent	NO	YES	NO	YES	NO	YES
Lagged inputs	NO	NO	NO	NO	YES	YES
Fixed effects	NO	NO	YES	YES	NO	NO

Notes: Each specification includes controls as the age of the primary caregiver, the number of siblings, the log of family income, and dummies for the college education of the primary caregiver, the presence of both biological parents at home, the gender of the study child, the indigenous status of the study child, the study child's cohort, and the day of data collection.

Table B.13: Estimated parameters of production function without parenting style for non-cognitive skills at age 8-9

	OLS	VA	FE	AB	CU	CV
Educational time parents	0.004	0.002	0.000	0.003	0.003	0.004
	(0.004)	(0.003)	(0.003)	(0.002)	(0.004)	(0.003)
Educational time others	0.000	0.007	-0.013	-0.002	0.008	-0.000
	(0.016)	(0.014)	(0.011)	(0.012)	(0.016)	(0.012)
Care time parents	-0.003	-0.001	-0.001	-0.001	-0.003	-0.001
	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
Care time others	0.003	-0.002	0.000	-0.002	0.003	-0.002
	(0.007)	(0.004)	(0.004)	(0.004)	(0.008)	(0.004)
Lagged test outcome		0.717^{***}		0.202^{***}		0.712^{***}
		(0.018)		(0.019)		(0.017)
Observations	2,876	2,759	6,605	6,508	2,606	2,570
Lagged dependent	NO	YES	NO	YES	NO	YES
Lagged inputs	NO	NO	NO	NO	YES	YES
Fixed effects	NO	NO	YES	YES	NO	NO

Notes: Each specification includes controls as the age of the primary caregiver, the number of siblings, the log of family income, and dummies for the college education of the primary caregiver, the presence of both biological parents at home, the gender of the study child, the indigenous status of the study child, the study child's cohort, and the day of data collection.

Table B.14: Estimated parameters of production function without parenting style for non-cognitive skills at age 10-11

	OLS	VA	FE	AB	CU	CV
Educational time parents	0.000	-0.001	0.000	-0.000	-0.001	-0.000
	(0.003)	(0.002)	(0.002)	(0.002)	(0.007)	(0.003)
Educational time others	-0.003	-0.001	-0.004	0.002	0.014	0.008
	(0.006)	(0.004)	(0.005)	(0.005)	(0.012)	(0.008)
Care time parents	-0.001	-0.000	-0.001	-0.001	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
Care time others	0.005^{**}	0.003^{*}	-0.000	0.001	-0.002	0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.004)	(0.003)
Lagged test outcome		0.731^{***}		0.202^{***}		0.735^{***}
		(0.012)		(0.019)		(0.018)
Observations	7,328	6,728	6,605	6,508	2,454	2,441
Lagged dependent	NO	YES	NO	YES	NO	YES
Lagged inputs	NO	NO	NO	NO	YES	YES
Fixed effects	NO	NO	YES	YES	NO	NO

Notes: Each specification includes controls as the age of the primary caregiver, the number of siblings, the log of family income, and dummies for the college education of the primary caregiver, the presence of both biological parents at home, the gender of the study child, the indigenous status of the study child, the study child's cohort, and the day of data collection.

Table B.15: Estimated parameters of production function without parenting style for non-cognitive skills at age 12-13

	OLS	VA	FE	AB	CU	CV
Educational time parents	0.001	0.002	0.001	0.001	0.001	0.003
	(0.003)	(0.002)	(0.002)	(0.002)	(0.005)	(0.004)
Educational time others	0.001	0.001	0.001	-0.002	-0.020	-0.007
	(0.007)	(0.004)	(0.005)	(0.004)	(0.018)	(0.013)
Care time parents	0.003**	0.001	0.000	-0.000	0.004*	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
Care time others	0.002	-0.000	0.001	0.001	0.006**	0.002
	(0.002)	(0.002)	(0.001)	(0.001)	(0.003)	(0.002)
Lagged test outcome		0.762***		0.202***		0.740***
		(0.012)		(0.019)		(0.021)
Observations	6,574	6,371	6,605	6,508	2,237	2,233
Lagged dependent	NO	YES	NO	YES	NO	YES
Lagged inputs	NO	NO	NO	NO	YES	YES
Fixed effects	NO	NO	YES	YES	NO	NO

Notes: Each specification includes controls as the age of the primary caregiver, the number of siblings, the log of family income, and dummies for the college education of the primary caregiver, the presence of both biological parents at home, the gender of the study child, the indigenous status of the study child, the study child's cohort, and the day of data collection.

Table B.16: Estimated parameters of production function without parenting style for non-cognitive skills at age 14-15

	OLS	VA	FE	AB	CU	CV
Educational time parents	0.002	0.002	0.001	0.003	0.007*	0.002
	(0.003)	(0.002)	(0.002)	(0.002)	(0.004)	(0.002)
Educational time others	0.003	0.000	-0.002	-0.001	0.002	0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)
Care time parents	0.002	0.000	-0.000	-0.000	0.004**	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
Care time others	-0.002	-0.002	-0.001	-0.001	-0.002	-0.002
	(0.003)	(0.002)	(0.002)	(0.002)	(0.004)	(0.003)
Lagged test outcome		0.748***		0.202***		0.731***
		(0.014)		(0.019)		(0.022)
Observations	5,765	5,564	6,605	6,508	1,905	1,897
Lagged dependent	NO	YES	NO	YES	NO	YES
Lagged inputs	NO	NO	NO	NO	YES	YES
Fixed effects	NO	NO	YES	YES	NO	NO

Notes: Each specification includes controls as the age of the primary caregiver, the number of siblings, the log of family income, and dummies for the college education of the primary caregiver, the presence of both biological parents at home, the gender of the study child, the indigenous status of the study child, the study child's cohort, and the day of data collection.

Table B.17: Estimated parameters of production function with aggregated parenting dimensions for non-cognitive skills at age 8-9

	OLS	VA	FE	AB	CU	CV
Emphatic style	0.011	-0.014	0.021	0.004	0.010	0.019
	(0.020)	(0.014)	(0.016)	(0.013)	(0.022)	(0.016)
Harsh style	-0.450***	-0.195***	-0.205***	-0.138***	-0.363***	-0.231***
	(0.019)	(0.016)	(0.018)	(0.015)	(0.024)	(0.019)
Educational time parents	-0.001	0.000	0.001	0.004	-0.001	0.002
	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)	(0.003)
Educational time others	0.005	-0.001	-0.013	-0.012	0.009	-0.001
	(0.014)	(0.012)	(0.010)	(0.010)	(0.015)	(0.012)
Care time parents	-0.003	-0.001	-0.001	-0.000	-0.002	-0.001
	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
Care time others	0.005	-0.001	0.000	-0.003	0.004	-0.001
	(0.007)	(0.004)	(0.004)	(0.004)	(0.008)	(0.005)
Lagged test outcome		0.645***		0.252***		0.640***
		(0.020)		(0.019)		(0.020)
Observations	2,780	2,667	6,599	6,463	2,419	2,417
Lagged dependent	NO	YES	NO	YES	NO	YES
Lagged inputs	NO	NO	NO	NO	YES	YES
Fixed effects	NO	NO	YES	YES	NO	NO

Notes: We define emphatic parenting style as one that is high on warmth, reasoning, and attempted consistency, while harsh parenting is defined as one that is high on hostility and inconsistency. Each specification includes controls as the age of the primary caregiver, the number of siblings, the log of family income, and dummies for the college education of the primary caregiver, the presence of both biological parents at home, the gender of the study child, the indigenous status of the study child, the study child's cohort, and the day of data collection.

Table B.18: Estimated parameters of production function with aggregated parenting dimensions for non-cognitive skills at age 10-11

	OLS	VA	FE	AB	CU	CV
Emphatic style	-0.020*	-0.012	0.006	-0.001	-0.026	-0.013
	(0.011)	(0.009)	(0.011)	(0.010)	(0.022)	(0.019)
Harsh style	-0.467***	-0.209***	-0.202***	-0.211***	-0.376***	-0.281***
	(0.013)	(0.011)	(0.013)	(0.013)	(0.029)	(0.026)
Educational time parents	-0.002	-0.002	-0.000	-0.000	-0.002	0.001
	(0.003)	(0.002)	(0.002)	(0.002)	(0.005)	(0.003)
Educational time others	-0.000	0.000	-0.003	0.002	0.013	0.010
	(0.005)	(0.004)	(0.005)	(0.004)	(0.008)	(0.007)
Care time parents	-0.002	-0.001	-0.000	-0.001	0.000	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
Care time others	0.007***	0.004**	0.001	0.002	-0.001	0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)
Lagged test outcome		0.646***		0.252***		0.652***
		(0.013)		(0.019)		(0.021)
Observations	7,299	6,703	6,599	6,463	2,267	2,264
Lagged dependent	NO	YES	NO	YES	NO	YES
Lagged inputs	NO	NO	NO	NO	YES	YES
Fixed effects	NO	NO	YES	YES	NO	NO

Notes: We define emphatic parenting style as one that is high on warmth, reasoning, and attempted consistency, while harsh parenting is defined as one that is high on hostility and inconsistency. Each specification includes controls as the age of the primary caregiver, the number of siblings, the log of family income, and dummies for the college education of the primary caregiver, the presence of both biological parents at home, the gender of the study child, the indigenous status of the study child, the study child's cohort, and the day of data collection.

Table B.19: Estimated parameters of production function with aggregated parenting dimensions for non-cognitive skills at age 12-13

	OLS	VA	FE	AB	CU	CV
Emphatic style	-0.065***	-0.027***	-0.024**	-0.009	-0.032	-0.006
	(0.013)	(0.009)	(0.010)	(0.009)	(0.023)	(0.018)
Harsh style	-0.465***	-0.217***	-0.220***	-0.228***	-0.372***	-0.315***
	(0.014)	(0.011)	(0.012)	(0.013)	(0.030)	(0.027)
Educational time parents	-0.003	0.000	0.000	0.001	-0.008*	-0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.004)	(0.003)
Educational time others	-0.008	-0.004	-0.004	-0.005	-0.010	-0.010
	(0.005)	(0.004)	(0.004)	(0.004)	(0.011)	(0.008)
Care time parents	0.001	0.000	0.000	-0.000	0.000	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
Care time others	0.000	-0.001	0.001	0.000	0.004	0.000
	(0.002)	(0.001)	(0.001)	(0.001)	(0.003)	(0.002)
Lagged test outcome		0.669***		0.252***		0.662***
		(0.013)		(0.019)		(0.022)
Observations	6,544	6,346	6,599	6,463	2,067	2,066
Lagged dependent	NO	YES	NO	YES	NO	YES
Lagged inputs	NO	NO	NO	NO	YES	YES
Fixed effects	NO	NO	YES	YES	NO	NO

Notes: We define emphatic parenting style as one that is high on warmth, reasoning, and attempted consistency, while harsh parenting is defined as one that is high on hostility and inconsistency. Each specification includes controls as the age of the primary caregiver, the number of siblings, the log of family income, and dummies for the college education of the primary caregiver, the presence of both biological parents at home, the gender of the study child, the indigenous status of the study child, the study child's cohort, and the day of data collection.

Table B.20: Estimated parameters of production function with aggregated parenting dimensions for non-cognitive skills at age 14-15

	OLS	VA	FE	AB	CU	CV
Emphatic style	-0.088***	-0.048***	-0.026**	-0.018*	-0.075***	-0.068***
	(0.013)	(0.010)	(0.011)	(0.011)	(0.025)	(0.020)
Harsh style	-0.485***	-0.238***	-0.243***	-0.220***	-0.372***	-0.309***
	(0.015)	(0.013)	(0.014)	(0.014)	(0.039)	(0.031)
Educational time parents	0.002	0.002	0.002	0.002	0.005*	0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)
Educational time others	0.002	0.000	-0.002	-0.002	0.002	0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.004)	(0.003)
Care time parents	-0.001	-0.001	-0.001	-0.001	0.001	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
Care time others	-0.001	-0.001	-0.001	-0.001	-0.002	-0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)
Lagged test outcome		0.644***		0.252***		0.655***
		(0.014)		(0.019)		(0.023)
Observations	5,726	5,531	6,599	6,463	1,753	1,753
Lagged dependent	NO	YES	NO	YES	NO	YES
Lagged inputs	NO	NO	NO	NO	YES	YES
Fixed effects	NO	NO	YES	YES	NO	NO

Notes: We define emphatic parenting style as one that is high on warmth, reasoning, and attempted consistency, while harsh parenting is defined as one that is high on hostility and inconsistency. Each specification includes controls as the age of the primary caregiver, the number of siblings, the log of family income, and dummies for the college education of the primary caregiver, the presence of both biological parents at home, the gender of the study child, the indigenous status of the study child, the study child's cohort, and the day of data collection.

Table B.21: Estimated parameters of production function with interaction between parenting styles and time investments for non-cognitive skills at age 8-9

	OLS	VA	FE	AB	CU	CV
Educational time:						
parents	-0.007	-0.004	0.001	0.005	-0.002	0.001
	(0.006)	(0.005)	(0.005)	(0.004)	(0.005)	(0.004)
parents x high emphatic style	0.012	0.006	0.000	-0.002	0.005	0.003
	(0.007)	(0.006)	(0.005)	(0.005)	(0.007)	(0.006)
parents x high harsh style	0.005	0.005	-0.000	-0.000	0.001	0.001
	(0.007)	(0.006)	(0.005)	(0.005)	(0.007)	(0.006)
others	0.018	-0.017	-0.003	0.001	0.026	0.004
	(0.036)	(0.032)	(0.028)	(0.030)	(0.040)	(0.030)
others x high emphatic style	-0.011	0.009	-0.018	-0.017	-0.011	-0.014
	(0.036)	(0.031)	(0.027)	(0.029)	(0.040)	(0.029)
others x high harsh style	-0.002	0.021	0.008	-0.002	-0.013	0.004
	(0.033)	(0.030)	(0.027)	(0.026)	(0.039)	(0.027)
Care time:						
parents	0.000	0.003	0.001	-0.000	-0.002	-0.000
	(0.003)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)
parents x high emphatic style	0.002	-0.002	-0.003	-0.001	0.004	0.000
	(0.004)	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)
parents x high harsh style	-0.008**	-0.005*	-0.002	-0.000	-0.004	-0.002
	(0.004)	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)
others	0.013	0.001	0.003	0.001	0.023*	0.007
	(0.012)	(0.007)	(0.007)	(0.006)	(0.012)	(0.007)
others x high emphatic style	-0.027	-0.010	-0.003	-0.001	-0.036	-0.013
	(0.022)	(0.011)	(0.009)	(0.008)	(0.023)	(0.010)
others x high harsh style	0.000	0.001	-0.004	-0.005	-0.008	-0.004
	(0.016)	(0.008)	(0.008)	(0.007)	(0.015)	(0.008)
Observations	2,780	2,667	6,599	6,463	2,419	2,417
Lagged dependent	NO	YES	NO	YES	NO	YES
Lagged inputs	NO	NO	NO	NO	YES	YES
Fixed effects	NO	NO	YES	YES	NO	NO

Notes: Emphatic parenting style is defined as high on warmth, reasoning, and attempted consistency. Harsh parenting is defined as high on hostility and inconsistency. In the estimation, we use dummies for high emphatic style and high harsh style if the corresponding parenting style is above median in the age group. Each specification includes controls as the age of the primary caregiver, the number of siblings, the log of family income, and dummies for the college education of the primary caregiver, the presence of both biological parents at home, the gender of the study child, the indigenous status of the study child, the study child's cohort, and the day of data collection.

Table B.22: Estimated parameters of production function with interaction between parenting styles and time investments for non-cognitive skills at age 10-11

	OLS	VA	FE	AB	CU	CV
Educational time:						
parents	-0.001	-0.000	-0.002	0.001	0.012	0.006
	(0.004)	(0.003)	(0.004)	(0.003)	(0.011)	(0.006)
parents x high emphatic style	-0.005	-0.003	0.001	-0.003	-0.019	-0.006
	(0.006)	(0.004)	(0.004)	(0.004)	(0.013)	(0.006)
parents x high harsh style	0.006	0.001	0.003	0.001	-0.010	-0.005
	(0.006)	(0.004)	(0.004)	(0.004)	(0.013)	(0.006)
others	0.002	-0.002	-0.008	-0.001	-0.035*	-0.030**
	(0.006)	(0.004)	(0.009)	(0.009)	(0.020)	(0.013)
others x high emphatic style	-0.001	0.001	0.009	0.005	0.061***	0.050***
	(0.011)	(0.008)	(0.011)	(0.010)	(0.016)	(0.012)
others x high harsh style	-0.004	0.004	0.001	0.003	0.030	0.023*
	(0.010)	(0.007)	(0.009)	(0.010)	(0.018)	(0.012)
Care time:						
parents	0.000	0.001	-0.002	-0.001	0.003	0.002
	(0.002)	(0.001)	(0.001)	(0.001)	(0.003)	(0.002)
parents x high emphatic style	-0.002	-0.000	0.002	0.001	-0.001	-0.001
	(0.002)	(0.001)	(0.002)	(0.002)	(0.003)	(0.003)
parents x high harsh style	-0.002	-0.002	0.001	-0.001	-0.005	-0.005*
	(0.002)	(0.001)	(0.002)	(0.002)	(0.003)	(0.003)
others	0.002	0.000	-0.002	-0.001	-0.000	0.001
	(0.003)	(0.003)	(0.002)	(0.003)	(0.005)	(0.004)
others x high emphatic style	0.003	-0.001	0.003	-0.000	-0.007	-0.006
	(0.004)	(0.003)	(0.003)	(0.003)	(0.007)	(0.005)
others x high harsh style	0.006	0.007**	0.001	0.005	0.002	0.006
	(0.004)	(0.003)	(0.003)	(0.003)	(0.007)	(0.005)
Observations	7,299	6,703	6,599	6,463	2,267	2,264
Lagged dependent	NO	YES	NO	YES	NO	YES
Lagged inputs	NO	NO	NO	NO	YES	YES
Fixed effects	NO	NO	YES	YES	NO	NO

Notes: Emphatic parenting style is defined as high on warmth, reasoning, and attempted consistency. Harsh parenting is defined as high on hostility and inconsistency. In the estimation, we use dummies for high emphatic style and high harsh style if the corresponding parenting style is above median in the age group. Each specification includes controls as the age of the primary caregiver, the number of siblings, the log of family income, and dummies for the college education of the primary caregiver, the presence of both biological parents at home, the gender of the study child, the indigenous status of the study child, the study child's cohort, and the day of data collection.

Table B.23: Estimated parameters of production function with interaction between parenting styles and time investments for non-cognitive skills at age 12-13

	OLS	VA	FE	AB	CU	CV
Educational time:						
parents	0.003	0.006**	0.002	0.002	0.005	0.005
	(0.004)	(0.003)	(0.002)	(0.002)	(0.006)	(0.005)
parents x high emphatic style	-0.005	-0.006*	-0.004	-0.004	-0.001	-0.003
	(0.005)	(0.003)	(0.004)	(0.003)	(0.009)	(0.007)
parents x high harsh style	-0.005	-0.004	0.003	0.004	-0.026***	-0.006
	(0.005)	(0.004)	(0.004)	(0.003)	(0.009)	(0.007)
others	0.001	0.005	-0.002	-0.000	0.031	0.039***
	(0.010)	(0.007)	(0.006)	(0.006)	(0.020)	(0.011)
others x high emphatic style	-0.002	-0.007	0.003	-0.005	-0.076**	-0.067**
	(0.012)	(0.008)	(0.009)	(0.008)	(0.039)	(0.028)
others x high harsh style	-0.011	-0.011	-0.005	0.000	-0.054**	-0.053***
	(0.014)	(0.011)	(0.009)	(0.009)	(0.026)	(0.016)
Care time:						
parents	0.001	0.002	0.001	-0.000	-0.002	-0.000
	(0.002)	(0.001)	(0.001)	(0.001)	(0.003)	(0.002)
parents x high emphatic style	0.001	-0.001	-0.001	0.000	0.004	-0.000
	(0.002)	(0.002)	(0.002)	(0.002)	(0.004)	(0.003)
parents x high harsh style	0.001	-0.001	0.000	-0.000	0.004	0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.004)	(0.003)
others	-0.000	-0.002	0.001	-0.001	0.008**	0.000
	(0.003)	(0.002)	(0.002)	(0.002)	(0.004)	(0.003)
others x high emphatic style	0.006	0.004	0.001	0.004	-0.001	0.005
	(0.004)	(0.003)	(0.003)	(0.003)	(0.005)	(0.004)
others x high harsh style	-0.002	0.000	0.000	0.000	-0.007	-0.002
	(0.004)	(0.003)	(0.003)	(0.003)	(0.005)	(0.004)
Observations	6,544	6,346	6,599	6,463	2,067	2,066
Lagged dependent	NO	YES	NO	YES	NO	YES
Lagged inputs	NO	NO	NO	NO	YES	YES
Fixed effects	NO	NO	YES	YES	NO	NO

Notes: Emphatic parenting style is defined as high on warmth, reasoning, and attempted consistency. Harsh parenting is defined as high on hostility and inconsistency. In the estimation, we use dummies for high emphatic style and high harsh style if the corresponding parenting style is above median in the age group. Each specification includes controls as the age of the primary caregiver, the number of siblings, the log of family income, and dummies for the college education of the primary caregiver, the presence of both biological parents at home, the gender of the study child, the indigenous status of the study child, the study child's cohort, and the day of data collection.

Table B.24: Estimated parameters of production function with interaction between parenting styles and time investments for non-cognitive skills at age 14-15

	OLS	VA	FE	AB	CU	CV
Educational time:						
parents	0.006**	0.004*	0.004*	0.003	-0.000	0.000
	(0.003)	(0.002)	(0.002)	(0.003)	(0.004)	(0.003)
parents x high emphatic style	-0.003	-0.002	-0.004	-0.001	0.014**	0.006
	(0.005)	(0.003)	(0.003)	(0.003)	(0.006)	(0.004)
parents x high harsh style	-0.004	-0.002	0.001	0.000	0.002	-0.003
	(0.005)	(0.003)	(0.003)	(0.003)	(0.007)	(0.005)
others	0.005*	0.004*	0.003	0.004	0.001	0.007
	(0.003)	(0.002)	(0.002)	(0.003)	(0.005)	(0.005)
others x high emphatic style	-0.008*	-0.007**	-0.007*	-0.007**	-0.005	-0.011*
	(0.004)	(0.003)	(0.004)	(0.003)	(0.007)	(0.006)
others x high harsh style	0.001	-0.002	-0.003	-0.005	0.005	-0.003
	(0.004)	(0.003)	(0.003)	(0.003)	(0.006)	(0.006)
Care time:						
parents	0.001	0.000	-0.001	0.001	-0.003	-0.000
	(0.002)	(0.002)	(0.001)	(0.002)	(0.003)	(0.002)
parents x high emphatic style	-0.001	-0.001	-0.001	-0.000	0.006	0.002
	(0.003)	(0.002)	(0.002)	(0.002)	(0.004)	(0.003)
parents x high harsh style	0.000	-0.000	0.002	-0.002	0.006	-0.000
	(0.003)	(0.002)	(0.002)	(0.002)	(0.004)	(0.003)
others	-0.006	-0.005*	-0.001	-0.007**	0.004	-0.003
	(0.004)	(0.003)	(0.003)	(0.003)	(0.007)	(0.005)
others x high emphatic style	-0.000	0.003	-0.000	0.008**	-0.002	0.006
	(0.005)	(0.004)	(0.004)	(0.004)	(0.007)	(0.006)
others x high harsh style	0.011**	0.004	0.001	0.003	-0.004	-0.004
	(0.005)	(0.004)	(0.004)	(0.004)	(0.007)	(0.005)
Observations	5,726	5,531	6,599	6,463	1,753	1,753
Lagged dependent	NO	YES	NO	YES	NO	YES
Lagged inputs	NO	NO	NO	NO	YES	YES
Fixed effects	NO	NO	YES	YES	NO	NO

Notes: Emphatic parenting style is defined as high on warmth, reasoning, and attempted consistency. Harsh parenting is defined as high on hostility and inconsistency. In the estimation, we use dummies for high emphatic style and high harsh style if the corresponding parenting style is above median in the age group. Each specification includes controls as the age of the primary caregiver, the number of siblings, the log of family income, and dummies for the college education of the primary caregiver, the presence of both biological parents at home, the gender of the study child, the indigenous status of the study child, the study child's cohort, and the day of data collection.

Table B.25: Estimated parameters of production function with non-linear parenting styles for non-cognitive skills at age 8-9

	Quartile		
	2nd	3rd	4th
Parental warmth	0.068*	0.014	0.012
	(0.035)	(0.034)	(0.043)
Parental reasoning	-0.040	-0.004	-0.040
	(0.034)	(0.034)	(0.044)
Hostile parenting	-0.097***	-0.121***	-0.268***
	(0.033)	(0.032)	(0.041)
Inconsistent parenting	-0.029	-0.021	-0.083**
	(0.034)	(0.034)	(0.039)
Attempted consistency	-0.032	-0.036	-0.060
	(0.036)	(0.036)	(0.037)
Observations		6,463	

Notes: We use dummy variables to indicate whether a child belongs to the 2nd, 3rd, or 4th quartile of parenting dimensions distributions (with the 1st quartile as the baseline category). Each specification includes the child's fixed effect, the lagged value of non-cognitive skills, time investments such as educational time with parents, educational time with other adults, care time with parents, and care time with other adults, as well as control variables such as the age of the primary caregiver, the number of siblings, the logarithm of family income, and dummy variables for the primary caregiver's college education, the presence of both biological parents at home, the gender of the study child, the study child's cohort, and the day of data collection.

Table B.26: Estimated parameters of production function with non-linear parenting styles for non-cognitive skills at age 10-11

	Quartile		
	2nd	3rd	4th
Parental warmth	0.013	0.028	0.043
	(0.027)	(0.026)	(0.030)
Parental reasoning	-0.021	-0.069***	-0.067**
	(0.024)	(0.026)	(0.030)
Hostile parenting	-0.114***	-0.164***	-0.330***
	(0.025)	(0.024)	(0.031)
Inconsistent parenting	-0.049*	-0.119***	-0.167***
	(0.026)	(0.026)	(0.028)
Attempted consistency	-0.034	0.020	-0.026
	(0.026)	(0.027)	(0.028)
Observations		6,463	

Notes: We use dummy variables to indicate whether a child belongs to the 2nd, 3rd, or 4th quartile of parenting dimensions distributions (with the 1st quartile as the baseline category). Each specification includes the child's fixed effect, the lagged value of non-cognitive skills, time investments such as educational time with parents, educational time with other adults, care time with parents, and care time with other adults, as well as control variables such as the age of the primary caregiver, the number of siblings, the logarithm of family income, and dummy variables for the primary caregiver's college education, the presence of both biological parents at home, the gender of the study child, the study child's cohort, and the day of data collection.

Table B.27: Estimated parameters of production function with non-linear parenting styles for non-cognitive skills at age 12-13

	Quartile		
	2nd	3rd	4th
Parental warmth	0.085***	0.094***	0.123***
	(0.025)	(0.028)	(0.031)
Parental reasoning	-0.055**	-0.078***	-0.073***
	(0.027)	(0.022)	(0.027)
Hostile parenting	-0.035	-0.171***	-0.337***
	(0.022)	(0.025)	(0.032)
Inconsistent parenting	-0.054**	-0.066***	-0.151***
	(0.024)	(0.024)	(0.028)
Attempted consistency	0.025	0.003	0.007
	(0.024)	(0.024)	(0.026)
Observations		6,463	

Notes: We use dummy variables to indicate whether a child belongs to the 2nd, 3rd, or 4th quartile of parenting dimensions distributions (with the 1st quartile as the baseline category). Each specification includes the child's fixed effect, the lagged value of non-cognitive skills, time investments such as educational time with parents, educational time with other adults, care time with parents, and care time with other adults, as well as control variables such as the age of the primary caregiver, the number of siblings, the logarithm of family income, and dummy variables for the primary caregiver's college education, the presence of both biological parents at home, the gender of the study child, the study child's cohort, and the day of data collection.

Table B.28: Estimated parameters of production function with non-linear parenting styles, lagged for non-cognitive skills at age 14-15

	Quartile		
	2nd	3rd	4th
Parental warmth	0.026	0.018	0.103***
	(0.031)	(0.031)	(0.036)
Parental reasoning	-0.047*	-0.087**	-0.097***
	(0.026)	(0.034)	(0.032)
Hostile parenting	-0.056**	-0.114***	-0.321***
	(0.027)	(0.029)	(0.036)
Inconsistent parenting	0.023	-0.024	-0.160***
	(0.029)	(0.029)	(0.034)
Attempted consistency	-0.012	0.059**	0.030
	(0.029)	(0.030)	(0.031)
Observations		6,463	

Notes: We use dummy variables to indicate whether a child belongs to the 2nd, 3rd, or 4th quartile of parenting dimensions distributions (with the 1st quartile as the baseline category). Each specification includes the child's fixed effect, the lagged value of non-cognitive skills, time investments such as educational time with parents, educational time with other adults, care time with parents, and care time with other adults, as well as control variables such as the age of the primary caregiver, the number of siblings, the logarithm of family income, and dummy variables for the primary caregiver's college education, the presence of both biological parents at home, the gender of the study child, the study child's cohort, and the day of data collection.

B.1.2. Cognitive skills

Table B.29: Estimated parameters of production function for cognitive skills (MRT) at age 10-11

	OLS	VA	FE	AB	CU	CV
Parental warmth	-0.025	-0.010	-0.024	-0.018	-0.014	-0.007
	(0.017)	(0.014)	(0.022)	(0.031)	(0.036)	(0.029)
Parental reasoning	0.000	0.005	-0.006	-0.019	-0.046*	-0.033
	(0.015)	(0.013)	(0.021)	(0.028)	(0.028)	(0.025)
Hostile parenting	0.001	0.009	0.012	0.010	0.010	0.011
	(0.016)	(0.014)	(0.022)	(0.030)	(0.031)	(0.026)
Inconsistent parenting	-0.082***	-0.052***	-0.029	0.003	-0.053	-0.036
	(0.016)	(0.013)	(0.020)	(0.030)	(0.032)	(0.026)
Attempted consistency	0.017	0.006	0.012	-0.001	0.025	0.025
	(0.014)	(0.012)	(0.020)	(0.026)	(0.027)	(0.023)
Educational time parents	-0.001	-0.002	-0.002	0.002	0.001	0.002
	(0.003)	(0.002)	(0.004)	(0.005)	(0.005)	(0.004)
Educational time others	0.003	0.002	-0.009	-0.001	0.005	-0.001
	(0.005)	(0.004)	(0.008)	(0.012)	(0.008)	(0.005)
Care time parents	0.001	0.001	0.004**	0.003	0.001	0.002
	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Care time others	0.005**	0.005***	0.007**	0.003	0.003	0.004
	(0.002)	(0.002)	(0.003)	(0.004)	(0.004)	(0.003)
Lagged test outcome		0.496***		0.179***		0.464***
		(0.012)		(0.039)		(0.019)
Observations	7,266	7,055	7,428	2,504	2,262	2,256
Lagged dependent	NO	YES	NO	YES	NO	YES
Lagged inputs	NO	NO	NO	NO	YES	YES
Fixed effects	NO	NO	YES	YES	NO	NO

Notes: Cognitive skills are measured using Matrix Reasoning Test outcomes standardized by age group. Each specification includes controls as the age of the primary caregiver, the number of siblings, the log of family income, and dummies for the college education of the primary caregiver, the presence of both biological parents at home, the gender of the study child, the indigenous status of the study child, the study child's cohort, and the day of data collection.

Table B.30: Estimated parameters of production function without parenting style for cognitive skills (MRT) at age 8-9

	OLS	VA	FE	AB	CU	CV
Educational time parents	0.015***	0.007*	-0.001	0.002	0.015***	0.006
	(0.004)	(0.004)	(0.003)	(0.004)	(0.005)	(0.004)
Educational time others	0.024	0.019	0.015	0.013	0.005	0.003
	(0.017)	(0.016)	(0.015)	(0.016)	(0.017)	(0.015)
Care time parents	-0.000	0.000	0.002	0.003	-0.002	-0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Care time others	0.004	0.003	0.002	0.008	0.010	0.006
	(0.006)	(0.006)	(0.005)	(0.006)	(0.006)	(0.006)
Lagged test outcome		0.461***		0.183***		0.461***
		(0.019)		(0.039)		(0.020)
Observations	2,862	2,794	7,497	2,617	2,594	2,587
Lagged dependent	NO	YES	NO	YES	NO	YES
Lagged inputs	NO	NO	NO	NO	YES	YES
Fixed effects	NO	NO	YES	YES	NO	NO

Notes: Cognitive skills are measured using Matrix Reasoning Test outcomes standardized by age group. Each specification includes controls as the age of the primary caregiver, the number of siblings, the log of family income, and dummies for the college education of the primary caregiver, the presence of both biological parents at home, the gender of the study child, the indigenous status of the study child, the study child's cohort, and the day of data collection.

Table B.31: Estimated parameters of production function without parenting style for cognitive skills (MRT) at age 10-11

	OLS	VA	FE	AB	CU	CV
Educational time parents	0.001	-0.001	-0.002	0.002	0.001	0.002
	(0.003)	(0.002)	(0.004)	(0.004)	(0.005)	(0.004)
Educational time others	0.003	0.002	-0.007	0.001	0.009	0.003
	(0.005)	(0.004)	(0.009)	(0.012)	(0.008)	(0.007)
Care time parents	0.000	0.000	0.004**	0.004*	0.000	0.002
	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Care time others	0.004**	0.004**	0.007**	0.004	0.003	0.004
	(0.002)	(0.002)	(0.003)	(0.004)	(0.003)	(0.003)
Lagged test outcome		0.502***		0.183***		0.458***
		(0.012)		(0.039)		(0.019)
Observations	7,349	7,129	7,497	2,617	2,454	2,446
Lagged dependent	NO	YES	NO	YES	NO	YES
Lagged inputs	NO	NO	NO	NO	YES	YES
Fixed effects	NO	NO	YES	YES	NO	NO

Notes: Cognitive skills are measured using Matrix Reasoning Test outcomes standardized by age group. Each specification includes controls as the age of the primary caregiver, the number of siblings, the log of family income, and dummies for the college education of the primary caregiver, the presence of both biological parents at home, the gender of the study child, the indigenous status of the study child, the study child's cohort, and the day of data collection.

Table B.32: Estimated parameters of production function without parenting style for cognitive skills (PPVT) at age 8-9

	OLS	VA	FE	AB	CU	CV
Educational time parents	0.018***	0.010***	0.002	-0.004	0.018***	0.010***
	(0.004)	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)
Educational time others	0.010	-0.001	-0.017	-0.000	0.003	-0.007
	(0.015)	(0.015)	(0.023)	(0.024)	(0.016)	(0.016)
Care time parents	0.002	0.000	-0.001	-0.003	-0.000	-0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Care time others	-0.005	-0.005	-0.006	-0.002	0.000	-0.003
	(0.005)	(0.005)	(0.009)	(0.010)	(0.006)	(0.006)
Lagged test outcome		0.499***		0.122***		0.499***
		(0.019)		(0.040)		(0.020)
Observations	2,864	2,732	3,501	2,329	2,596	2,530
Lagged dependent	NO	YES	NO	YES	NO	YES
Lagged inputs	NO	NO	NO	NO	YES	YES
Fixed effects	NO	NO	YES	YES	NO	NO

Notes: Cognitive skills are measured using Peabody Picture Vocabulary Test (PPVT) outcomes standardized by age group. Each specification includes controls as the age of the primary caregiver, the number of siblings, the log of family income, and dummies for the college education of the primary caregiver, the presence of both biological parents at home, the gender of the study child, the indigenous status of the study child, the study child's cohort, and the day of data collection.

Table B.33: Estimated parameters of production function with interaction between parenting styles and time investments for cognitive skills (MRT) at age 8-9

	OLS	VA	FE	AB	CU	CV
Educational time:						
parents	0.010	0.005	0.001	0.007	0.007	0.003
	(0.007)	(0.006)	(0.006)	(0.007)	(0.007)	(0.007)
parents x high emphatic style	0.000	-0.005	-0.006	-0.003	0.012	0.004
	(0.009)	(0.008)	(0.007)	(0.008)	(0.010)	(0.008)
parents x high harsh style	0.009	0.008	0.003	-0.005	0.007	0.006
	(0.009)	(0.007)	(0.007)	(0.008)	(0.010)	(0.008)
others	0.030	0.021	-0.013	-0.033	0.055	0.043
	(0.038)	(0.036)	(0.037)	(0.039)	(0.039)	(0.033)
others x high emphatic style	0.004	0.008	0.050	0.052	-0.027	-0.014
	(0.040)	(0.039)	(0.038)	(0.039)	(0.040)	(0.036)
others x high harsh style	-0.021	-0.020	-0.012	-0.001	-0.078**	-0.068**
	(0.032)	(0.032)	(0.035)	(0.035)	(0.031)	(0.027)
Care time:						
parents	-0.003	-0.002	-0.001	0.004	-0.005	-0.004
	(0.004)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)
parents x high emphatic style	0.003	0.004	0.006	0.006	0.007	0.007*
	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.004)
parents x high harsh style	0.002	0.000	-0.001	-0.006	0.001	0.001
	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.004)
others	0.013	0.011	0.004	0.008	0.017	0.011
	(0.011)	(0.011)	(0.011)	(0.012)	(0.013)	(0.012)
others x high emphatic style	-0.006	-0.007	-0.017	0.000	-0.001	-0.011
	(0.016)	(0.016)	(0.016)	(0.015)	(0.018)	(0.016)
others x high harsh style	-0.009	-0.007	0.008	0.008	0.002	0.003
	(0.013)	(0.013)	(0.013)	(0.013)	(0.015)	(0.013)
Observations	2,753	2,690	7,428	2,504	2,399	2,392
Lagged dependent	NO	YES	NO	YES	NO	YES
Lagged inputs	NO	NO	NO	NO	YES	YES
Fixed effects	NO	NO	YES	YES	NO	NO

Notes: Emphatic parenting style is defined as high on warmth, reasoning, and attempted consistency. Harsh parenting is defined as high on hostility and inconsistency. In the estimation, we use dummies for high emphatic style and high harsh style if the corresponding parenting style is above median in the age group. Each specification includes controls as the age of the primary caregiver, the number of siblings, the log of family income, and dummies for the college education of the primary caregiver, the presence of both biological parents at home, the gender of the study child, the indigenous status of the study child, the study child's cohort, and the day of data collection.

Table B.34: Estimated parameters of production function with interaction between parenting styles and time investments for cognitive skills (MRT) at age 10-11

	OLS	VA	FE	AB	CU	CV
Educational time:						
parents	0.005	0.004	-0.002	0.001	0.015	0.011
	(0.005)	(0.004)	(0.007)	(0.009)	(0.010)	(0.008)
parents x high emphatic style	-0.009*	-0.005	0.004	0.004	-0.009	-0.004
	(0.006)	(0.005)	(0.007)	(0.009)	(0.010)	(0.008)
parents x high harsh style	-0.003	-0.006	-0.005	-0.005	-0.016*	-0.012
	(0.006)	(0.005)	(0.007)	(0.009)	(0.009)	(0.007)
others	-0.004	-0.005	-0.021	-0.017	0.001	-0.012
	(0.007)	(0.005)	(0.021)	(0.027)	(0.017)	(0.013)
others x high emphatic style	-0.007	0.001	0.017	0.010	-0.011	0.001
	(0.009)	(0.007)	(0.019)	(0.025)	(0.017)	(0.013)
others x high harsh style	0.018**	0.013*	0.007	0.020	0.012	0.014
	(0.009)	(0.007)	(0.019)	(0.024)	(0.016)	(0.013)
Care time:						
parents	-0.001	-0.001	0.007**	0.003	0.003	0.003
	(0.002)	(0.002)	(0.003)	(0.004)	(0.003)	(0.003)
parents x high emphatic style	0.003	0.002	-0.003	0.002	-0.001	-0.000
	(0.002)	(0.002)	(0.003)	(0.004)	(0.004)	(0.003)
parents x high harsh style	0.000	0.001	-0.004	-0.003	-0.004	-0.002
	(0.002)	(0.002)	(0.003)	(0.004)	(0.004)	(0.003)
others	0.004	0.004	0.006	0.004	-0.004	-0.001
	(0.004)	(0.003)	(0.006)	(0.007)	(0.006)	(0.006)
others x high emphatic style	0.001	-0.000	-0.002	0.003	0.008	0.006
	(0.004)	(0.004)	(0.007)	(0.008)	(0.008)	(0.006)
others x high harsh style	0.001	0.002	0.004	-0.005	0.006	0.003
	(0.004)	(0.004)	(0.007)	(0.008)	(0.008)	(0.006)
Observations	7,266	7,055	7,428	2,504	2,262	2,256
Lagged dependent	NO	YES	NO	YES	NO	YES
Lagged inputs	NO	NO	NO	NO	YES	YES
Fixed effects	NO	NO	YES	YES	NO	NO

Notes: Emphatic parenting style is defined as high on warmth, reasoning, and attempted consistency. Harsh parenting is defined as high on hostility and inconsistency. In the estimation, we use dummies for high emphatic style and high harsh style if the corresponding parenting style is above median in the age group. Each specification includes controls as the age of the primary caregiver, the number of siblings, the log of family income, and dummies for the college education of the primary caregiver, the presence of both biological parents at home, the gender of the study child, the indigenous status of the study child, the study child's cohort, and the day of data collection.

Table B.35: Estimated parameters of production function with interaction between parenting styles and time investments for cognitive skills (PPVT) at age 8-9

	OLS	VA	FE	AB	CU	CV
Educational time:						
parents	0.017**	0.008	0.002	-0.003	0.018**	0.008
	(0.007)	(0.006)	(0.008)	(0.009)	(0.007)	(0.007)
parents x high emphatic style	-0.005	-0.001	0.002	-0.004	-0.004	0.000
	(0.008)	(0.007)	(0.009)	(0.009)	(0.009)	(0.008)
parents x high harsh style	0.004	0.005	-0.002	0.001	0.002	0.001
	(0.008)	(0.007)	(0.009)	(0.009)	(0.008)	(0.007)
others	0.003	-0.032	-0.073*	-0.084*	0.008	-0.031
	(0.034)	(0.029)	(0.041)	(0.043)	(0.038)	(0.029)
others x high emphatic style	0.026	0.072**	0.130***	0.150***	0.014	0.067**
	(0.032)	(0.029)	(0.042)	(0.045)	(0.036)	(0.028)
others x high harsh style	-0.009	-0.024	-0.046	-0.054	-0.033	-0.044
	(0.031)	(0.028)	(0.038)	(0.042)	(0.035)	(0.028)
Care time:						
parents	0.003	0.001	-0.001	-0.002	-0.001	-0.001
	(0.004)	(0.003)	(0.003)	(0.004)	(0.004)	(0.003)
parents x high emphatic style	0.000	-0.001	-0.002	-0.002	0.002	-0.000
	(0.004)	(0.004)	(0.004)	(0.005)	(0.004)	(0.004)
parents x high harsh style	-0.002	-0.001	0.001	0.001	-0.001	0.001
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
others	-0.005	-0.011	-0.016	-0.012	-0.001	-0.009
	(0.011)	(0.008)	(0.010)	(0.011)	(0.013)	(0.009)
others x high emphatic style	0.010	0.003	-0.012	-0.017	0.018	0.005
	(0.012)	(0.011)	(0.015)	(0.017)	(0.013)	(0.012)
others x high harsh style	0.002	0.015	0.031**	0.039**	0.009	0.022**
	(0.012)	(0.011)	(0.015)	(0.018)	(0.014)	(0.011)
Observations	2,755	2,633	3,437	2,156	2,401	2,343
Lagged dependent	NO	YES	NO	YES	NO	YES
Lagged inputs	NO	NO	NO	NO	YES	YES
Fixed effects	NO	NO	YES	YES	NO	NO

Notes: Emphatic parenting style is defined as high on warmth, reasoning, and attempted consistency. Harsh parenting is defined as high on hostility and inconsistency. In the estimation, we use dummies for high emphatic style and high harsh style if the corresponding parenting style is above median in the age group. Each specification includes controls as the age of the primary caregiver, the number of siblings, the log of family income, and dummies for the college education of the primary caregiver, the presence of both biological parents at home, the gender of the study child, the indigenous status of the study child, the study child's cohort, and the day of data collection.

Table B.36: Estimated parameters of production function with non-linear parenting styles for cognitive skills (MRT) at age 8-9

	Quartile		
	2nd	3rd	4th
Parental warmth	-0.066	-0.014	-0.109
	(0.057)	(0.064)	(0.079)
Parental reasoning	-0.058	0.011	0.038
	(0.057)	(0.062)	(0.076)
Hostile parenting	0.157**	0.093	-0.008
	(0.063)	(0.061)	(0.072)
Inconsistent parenting	-0.085	-0.007	-0.050
	(0.062)	(0.061)	(0.069)
Attempted consistency	-0.040	0.004	-0.039
	(0.060)	(0.059)	(0.067)
Observations		2,504	

Notes: We use dummy variables to indicate whether a child belongs to the 2nd, 3rd, or 4th quartile of parenting dimensions distributions (with the 1st quartile as the baseline category). Each specification includes the child's fixed effect, the lagged value of non-cognitive skills, time investments such as educational time with parents, educational time with other adults, care time with parents, and care time with other adults, as well as control variables such as the age of the primary caregiver, the number of siblings, the logarithm of family income, and dummy variables for the primary caregiver's college education, the presence of both biological parents at home, the gender of the study child, the study child's cohort, and the day of data collection.

Table B.37: Estimated parameters of production function with non-linear parenting styles for cognitive skills (MRT) at age 10-11

	Quartile		
	2nd	3rd	4th
Parental warmth	-0.142**	-0.051	-0.114
	(0.063)	(0.062)	(0.075)
Parental reasoning	0.006	-0.054	0.041
	(0.059)	(0.065)	(0.075)
Hostile parenting	0.126**	0.040	0.018
	(0.063)	(0.059)	(0.072)
Inconsistent parenting	-0.139**	-0.123**	-0.057
	(0.065)	(0.063)	(0.067)
Attempted consistency	0.046	0.122**	-0.021
	(0.062)	(0.062)	(0.068)
Observations		2,504	

Notes: We use dummy variables to indicate whether a child belongs to the 2nd, 3rd, or 4th quartile of parenting dimensions distributions (with the 1st quartile as the baseline category). Each specification includes the child's fixed effect, the lagged value of non-cognitive skills, time investments such as educational time with parents, educational time with other adults, care time with parents, and care time with other adults, as well as control variables such as the age of the primary caregiver, the number of siblings, the logarithm of family income, and dummy variables for the primary caregiver's college education, the presence of both biological parents at home, the gender of the study child, the study child's cohort, and the day of data collection.

Table B.38: Estimated parameters of production function with non-linear parenting styles for cognitive skills (MRT) at age 8-9

	Quartile		
	2nd	3rd	4th
Parental warmth	0.041	-0.025	-0.087
	(0.059)	(0.060)	(0.076)
Parental reasoning	0.055	-0.016	0.071
	(0.059)	(0.061)	(0.077)
Hostile parenting	-0.096	0.004	-0.091
	(0.062)	(0.061)	(0.067)
Inconsistent parenting	-0.102	0.005	0.051
	(0.064)	(0.060)	(0.064)
Attempted consistency	0.008	-0.036	0.005
	(0.062)	(0.063)	(0.064)
Observations		2,156	

Notes: We use dummy variables to indicate whether a child belongs to the 2nd, 3rd, or 4th quartile of parenting dimensions distributions (with the 1st quartile as the baseline category). Each specification includes the child's fixed effect, the lagged value of non-cognitive skills, time investments such as educational time with parents, educational time with other adults, care time with parents, and care time with other adults, as well as control variables such as the age of the primary caregiver, the number of siblings, the logarithm of family income, and dummy variables for the primary caregiver's college education, the presence of both biological parents at home, the gender of the study child, the study child's cohort, and the day of data collection.

B.2. Robustness checks

B.2.1. Non-cognitive skills

Table B.39: Estimated parameters of production function with jointly estimated parenting dimensions for non-cognitive skills at age 8-9

	OLS	VA	FE	AB	CU	CV
Warm style	0.153***	0.061***	0.085***	0.042***	0.128***	0.093***
	(0.018)	(0.013)	(0.017)	(0.013)	(0.022)	(0.018)
Reasoning style	-0.145***	-0.050***	-0.048***	-0.036**	-0.098***	-0.060***
	(0.019)	(0.015)	(0.016)	(0.015)	(0.024)	(0.019)
Hostile/inconstistent style	-0.413***	-0.192***	-0.194***	-0.135***	-0.339***	-0.221***
	(0.019)	(0.015)	(0.015)	(0.014)	(0.022)	(0.018)
Consistent style	0.072***	0.044***	0.019	0.015	0.052***	0.025*
	(0.016)	(0.012)	(0.012)	(0.012)	(0.018)	(0.014)
Educational time parents	-0.003	-0.000	0.001	0.003	-0.003	0.001
	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)	(0.003)
Educational time others	0.005	-0.000	-0.013	-0.012	0.009	0.000
	(0.015)	(0.012)	(0.010)	(0.010)	(0.015)	(0.012)
Care time parents	-0.002	-0.001	-0.001	-0.000	-0.001	-0.001
	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
Care time others	0.004	-0.001	0.000	-0.003	0.004	-0.002
	(0.007)	(0.004)	(0.004)	(0.004)	(0.008)	(0.005)
Lagged test outcome		0.633***		0.253***		0.631***
		(0.020)		(0.019)		(0.020)
Observations	2,780	2,667	6,599	6,463	2,419	2,417
Lagged dependent	NO	YES	NO	YES	NO	YES
Lagged inputs	NO	NO	NO	NO	YES	YES
Fixed effects	NO	NO	YES	YES	NO	NO

Notes: Each specification includes controls as the age of the primary caregiver, the number of siblings, the log of family income, and dummies for the college education of the primary caregiver, the presence of both biological parents at home, the gender of the study child, the indigenous status of the study child, the study child's cohort, and the day of data collection.

Table B.40: Estimated parameters of production function with jointly estimated parenting dimensions for non-cognitive skills at age 10-11

	OLS	VA	FE	AB	CU	CV
Warm style	0.162***	0.067***	0.082***	0.075***	0.123***	0.099***
	(0.012)	(0.009)	(0.011)	(0.011)	(0.025)	(0.021)
Reasoning style	-0.109***	-0.048***	-0.045***	-0.061***	-0.073***	-0.064***
	(0.011)	(0.009)	(0.010)	(0.010)	(0.024)	(0.019)
Hostile/inconstistent style	-0.433***	-0.201***	-0.189***	-0.194***	-0.361***	-0.266***
	(0.011)	(0.010)	(0.012)	(0.012)	(0.025)	(0.022)
Consistent style	0.082***	0.032***	0.029***	0.033***	0.066***	0.047***
	(0.010)	(0.008)	(0.009)	(0.009)	(0.020)	(0.016)
Educational time parents	-0.001	-0.002	-0.000	-0.001	-0.002	0.001
	(0.003)	(0.002)	(0.002)	(0.002)	(0.005)	(0.003)
Educational time others	-0.001	-0.000	-0.003	0.003	0.014*	0.011
	(0.005)	(0.004)	(0.005)	(0.004)	(0.008)	(0.007)
Care time parents	-0.002	-0.001	-0.000	-0.001	0.000	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
Care time others	0.007***	0.004**	0.001	0.002	-0.001	0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)
Lagged test outcome		0.632***		0.253***		0.641***
		(0.013)		(0.019)		(0.021)
Observations	7,299	6,703	6,599	6,463	2,267	2,264
Lagged dependent	NO	YES	NO	YES	NO	YES
Lagged inputs	NO	NO	NO	NO	YES	YES
Fixed effects	NO	NO	YES	YES	NO	NO

Notes: Each specification includes controls as the age of the primary caregiver, the number of siblings, the log of family income, and dummies for the college education of the primary caregiver, the presence of both biological parents at home, the gender of the study child, the indigenous status of the study child, the study child's cohort, and the day of data collection.

Table B.41: Estimated parameters of production function with jointly estimated parenting dimensions for non-cognitive skills at age 12-13

	OLS	VA	FE	AB	CU	CV
Warm style	0.188***	0.093***	0.087***	0.102***	0.186***	0.159***
	(0.013)	(0.009)	(0.011)	(0.011)	(0.027)	(0.022)
Reasoning style	-0.113***	-0.050***	-0.047***	-0.037***	-0.077***	-0.046***
	(0.011)	(0.008)	(0.008)	(0.008)	(0.021)	(0.016)
Hostile/inconstistent style	-0.424***	-0.200***	-0.207***	-0.208***	-0.317***	-0.273***
	(0.013)	(0.011)	(0.011)	(0.012)	(0.028)	(0.025)
Consistent style	0.041***	0.018**	0.020**	0.022**	0.016	0.018
	(0.013)	(0.009)	(0.010)	(0.010)	(0.025)	(0.019)
Educational time parents	-0.004*	-0.000	0.000	0.001	-0.008*	-0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.004)	(0.003)
Educational time others	-0.007	-0.004	-0.004	-0.005	-0.011	-0.010
	(0.005)	(0.004)	(0.004)	(0.004)	(0.011)	(0.009)
Care time parents	0.001	0.000	0.000	-0.000	0.000	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
Care time others	0.000	-0.001	0.001	0.000	0.004	0.000
	(0.002)	(0.001)	(0.001)	(0.001)	(0.003)	(0.002)
Lagged test outcome		0.658***		0.253***		0.653***
		(0.013)		(0.019)		(0.023)
Observations	6,544	6,346	6,599	6,463	2,067	2,066
Lagged dependent	NO	YES	NO	YES	NO	YES
Lagged inputs	NO	NO	NO	NO	YES	YES
Fixed effects	NO	NO	YES	YES	NO	NO

Notes: Each specification includes controls as the age of the primary caregiver, the number of siblings, the log of family income, and dummies for the college education of the primary caregiver, the presence of both biological parents at home, the gender of the study child, the indigenous status of the study child, the study child's cohort, and the day of data collection.

Table B.42: Estimated parameters of production function with jointly estimated parenting dimensions for non-cognitive skills at age 14-15

	OLS	VA	FE	AB	CU	CV
Warm style	0.197***	0.091***	0.095***	0.092***	0.096***	0.088***
	(0.014)	(0.010)	(0.012)	(0.012)	(0.034)	(0.025)
Reasoning style	-0.136***	-0.066***	-0.059***	-0.050***	-0.092***	-0.087***
	(0.011)	(0.009)	(0.010)	(0.010)	(0.021)	(0.018)
Hostile/inconstistent style	-0.434***	-0.220***	-0.233***	-0.207***	-0.349***	-0.293***
	(0.015)	(0.012)	(0.013)	(0.013)	(0.035)	(0.028)
Consistent style	-0.052***	-0.028***	-0.006	-0.011	-0.023	-0.038*
	(0.013)	(0.010)	(0.011)	(0.011)	(0.026)	(0.021)
Educational time parents	0.001	0.002	0.002	0.002	0.005*	0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)
Educational time others	0.002	0.000	-0.002	-0.002	0.003	0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.004)	(0.003)
Care time parents	-0.001	-0.001	-0.001	-0.001	0.001	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
Care time others	-0.001	-0.001	-0.001	-0.001	-0.003	-0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)
Lagged test outcome		0.635***		0.253***		0.651***
		(0.014)		(0.019)		(0.023)
Observations	5,726	5,531	6,599	6,463	1,753	1,753
Lagged dependent	NO	YES	NO	YES	NO	YES
Lagged inputs	NO	NO	NO	NO	YES	YES
Fixed effects	NO	NO	YES	YES	NO	NO

Notes: Each specification includes controls as the age of the primary caregiver, the number of siblings, the log of family income, and dummies for the college education of the primary caregiver, the presence of both biological parents at home, the gender of the study child, the indigenous status of the study child, the study child's cohort, and the day of data collection.

Table B.43: Estimated parameters of production function with extended set of controls for non-cognitive skills at age 8-9

	OLS	VA	FE	AB	CU	CV
Parental warmth	0.065***	0.021	0.049**	0.010	0.048*	0.041**
	(0.022)	(0.017)	(0.020)	(0.017)	(0.027)	(0.020)
Parental reasoning	-0.087***	-0.044***	-0.025*	-0.012	-0.055**	-0.031*
	(0.019)	(0.015)	(0.015)	(0.015)	(0.022)	(0.016)
Hostile parenting	-0.363***	-0.171***	-0.161***	-0.117***	-0.301***	-0.195***
	(0.023)	(0.018)	(0.016)	(0.017)	(0.027)	(0.020)
Inconsistent parenting	-0.147***	-0.059***	-0.062***	-0.045**	-0.101***	-0.070***
	(0.023)	(0.018)	(0.018)	(0.018)	(0.028)	(0.021)
Attempted consistency	0.004	-0.010	-0.016	-0.010	-0.011	-0.007
	(0.018)	(0.015)	(0.015)	(0.014)	(0.021)	(0.016)
Educational time parents	-0.002	-0.000	0.001	0.004	-0.003	0.001
	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)	(0.003)
Educational time others	0.010	0.004	-0.012	-0.012	0.016	0.005
	(0.015)	(0.012)	(0.010)	(0.010)	(0.015)	(0.012)
Care time parents	-0.002	-0.000	-0.001	-0.000	-0.001	-0.001
	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
Care time others	0.000	-0.004	-0.000	-0.003	0.000	-0.004
	(0.008)	(0.004)	(0.004)	(0.004)	(0.008)	(0.005)
Lagged test outcome		0.636***		0.254***		0.632***
		(0.020)		(0.019)		(0.020)
Observations	2,735	2,626	6,599	6,462	2,384	2,382
Lagged dependent	NO	YES	NO	YES	NO	YES
Lagged inputs	NO	NO	NO	NO	YES	YES
Fixed effects	NO	NO	YES	YES	NO	NO

Notes: Each specification includes controls as the age of the primary caregiver, the number of siblings, the log of family income, standardized Social-Economic Indexes for Areas, weight at birth and dummies for the college education of the primary caregiver, the presence of both biological parents at home, the gender of the study child, the indigenous status of th study child, the study child's cohort, the day of data collection, urban area, Australian state and early birth.

Table B.44: Estimated parameters of production function with extended set of controls for non-cognitive skills at age 10-11

	OLS	VA	FE	AB	CU	CV
Parental warmth	0.049***	0.011	0.037***	0.033**	0.041	0.042*
	(0.015)	(0.012)	(0.013)	(0.013)	(0.031)	(0.025)
Parental reasoning	-0.085***	-0.026***	-0.033***	-0.035***	-0.066***	-0.045**
	(0.012)	(0.009)	(0.010)	(0.010)	(0.023)	(0.018)
Hostile parenting	-0.386***	-0.179***	-0.170***	-0.160***	-0.308***	-0.216***
	(0.014)	(0.013)	(0.013)	(0.013)	(0.028)	(0.023)
Inconsistent parenting	-0.142***	-0.068***	-0.053***	-0.077***	-0.130***	-0.111***
	(0.014)	(0.011)	(0.013)	(0.013)	(0.030)	(0.023)
Attempted consistency	0.017	-0.001	0.006	0.005	-0.009	-0.007
	(0.011)	(0.009)	(0.010)	(0.010)	(0.022)	(0.017)
Educational time parents	-0.000	-0.001	-0.000	-0.001	-0.002	0.001
	(0.003)	(0.002)	(0.002)	(0.002)	(0.005)	(0.003)
Educational time others	-0.000	-0.000	-0.003	0.003	0.013	0.009
	(0.005)	(0.004)	(0.005)	(0.004)	(0.010)	(0.008)
Care time parents	-0.002*	-0.001	-0.000	-0.001	-0.000	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
Care time others	0.006***	0.004**	0.001	0.002	-0.002	0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)
Lagged test outcome		0.632***		0.254***		0.644***
		(0.013)		(0.019)		(0.021)
Observations	7,200	6,616	6,599	6,462	2,229	2,226
Lagged dependent	NO	YES	NO	YES	NO	YES
Lagged inputs	NO	NO	NO	NO	YES	YES
Fixed effects	NO	NO	YES	YES	NO	NO

Notes: Each specification includes controls as the age of the primary caregiver, the number of siblings, the log of family income, standardized Social-Economic Indexes for Areas, weight at birth and dummies for the college education of the primary caregiver, the presence of both biological parents at home, the gender of the study child, the indigenous status of th study child, the study child's cohort, the day of data collection, urban area, Australian state and early birth.

Table B.45: Estimated parameters of production function with extended set of controls for non-cognitive skills at age 12-13

	OLS	VA	FE	AB	CU	CV
Parental warmth	0.049***	0.025**	0.028**	0.044***	0.077**	0.058**
	(0.017)	(0.012)	(0.012)	(0.013)	(0.032)	(0.026)
Parental reasoning	-0.081***	-0.035***	-0.031***	-0.026***	-0.054**	-0.024
	(0.012)	(0.009)	(0.009)	(0.010)	(0.024)	(0.019)
Hostile parenting	-0.391***	-0.187***	-0.178***	-0.180***	-0.320***	-0.270***
	(0.017)	(0.013)	(0.013)	(0.014)	(0.030)	(0.026)
Inconsistent parenting	-0.115***	-0.052***	-0.065***	-0.059***	-0.049	-0.053**
	(0.017)	(0.012)	(0.013)	(0.012)	(0.031)	(0.024)
Attempted consistency	0.025*	0.011	0.002	0.005	0.005	0.007
	(0.013)	(0.009)	(0.010)	(0.010)	(0.021)	(0.017)
Educational time parents	-0.003	0.000	0.000	0.001	-0.008**	-0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.004)	(0.003)
Educational time others	-0.005	-0.003	-0.004	-0.005	-0.012	-0.013
	(0.005)	(0.004)	(0.004)	(0.004)	(0.012)	(0.009)
Care time parents	0.001	0.001	0.001	-0.000	0.001	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
Care time others	0.000	-0.001	0.001	0.000	0.003	0.000
	(0.002)	(0.001)	(0.001)	(0.001)	(0.003)	(0.002)
Lagged test outcome		0.654***		0.254***		0.644***
		(0.013)		(0.019)		(0.022)
Observations	6,476	6,283	6,599	6,462	2,042	2,041
Lagged dependent	NO	YES	NO	YES	NO	YES
Lagged inputs	NO	NO	NO	NO	YES	YES
Fixed effects	NO	NO	YES	YES	NO	NO

Notes: Each specification includes controls as the age of the primary caregiver, the number of siblings, the log of family income, standardized Social-Economic Indexes for Areas, weight at birth and dummies for the college education of the primary caregiver, the presence of both biological parents at home, the gender of the study child, the indigenous status of th study child, the study child's cohort, the day of data collection, urban area, Australian state and early birth.

Table B.46: Estimated parameters of production function with extended set of controls for non-cognitive skills at age 14-15

	OLS	VA	FE	AB	CU	CV
Parental warmth	0.053***	0.018	0.032**	0.036**	-0.020	0.000
	(0.017)	(0.013)	(0.014)	(0.014)	(0.039)	(0.028)
Parental reasoning	-0.087***	-0.037***	-0.040***	-0.033***	-0.040*	-0.043**
	(0.013)	(0.010)	(0.011)	(0.011)	(0.024)	(0.019)
Hostile parenting	-0.365***	-0.182***	-0.167***	-0.157***	-0.265***	-0.222***
	(0.019)	(0.015)	(0.016)	(0.015)	(0.036)	(0.028)
Inconsistent parenting	-0.151***	-0.085***	-0.103***	-0.080***	-0.168***	-0.139***
	(0.018)	(0.014)	(0.015)	(0.015)	(0.038)	(0.031)
Attempted consistency	0.023*	0.003	0.025**	0.016	0.016	-0.004
	(0.013)	(0.011)	(0.011)	(0.011)	(0.024)	(0.020)
Educational time parents	0.002	0.002	0.002	0.002	0.005*	0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)
Educational time others	0.003	0.001	-0.002	-0.002	0.003	0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)
Care time parents	-0.000	-0.001	-0.001	-0.001	0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
Care time others	-0.001	-0.001	-0.001	-0.001	-0.003	-0.003
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)
Lagged test outcome		0.631***		0.254***		0.644***
		(0.015)		(0.019)		(0.024)
Observations	5,671	5,478	6,599	6,462	1,733	1,733
Lagged dependent	NO	YES	NO	YES	NO	YES
Lagged inputs	NO	NO	NO	NO	YES	YES
Fixed effects	NO	NO	YES	YES	NO	NO

Notes: Each specification includes controls as the age of the primary caregiver, the number of siblings, the log of family income, standardized Social-Economic Indexes for Areas, weight at birth and dummies for the college education of the primary caregiver, the presence of both biological parents at home, the gender of the study child, the indigenous status of th study child, the study child's cohort, the day of data collection, urban area, Australian state and early birth.

B.2.2. Cognitive skills

Table B.47: Estimated parameters of production function with jointly estimated parenting dimensions for cognitive skills (MRT) at age 8-9

	OLS	VA	FE	AB	CU	CV
Warm style	-0.037*	-0.040**	-0.041**	-0.018	-0.040	-0.049*
	(0.022)	(0.019)	(0.019)	(0.028)	(0.030)	(0.026)
Reasoning style	-0.052**	-0.028	0.001	0.021	-0.011	0.009
	(0.022)	(0.019)	(0.020)	(0.029)	(0.028)	(0.024)
Hostile/inconstistent style	-0.027	-0.016	-0.008	-0.007	-0.010	-0.007
	(0.021)	(0.019)	(0.019)	(0.027)	(0.027)	(0.024)
Consistent style	0.017	0.020	0.024	0.020	0.036	0.040*
	(0.020)	(0.018)	(0.019)	(0.025)	(0.025)	(0.022)
Educational time parents	0.014***	0.006	-0.001	0.003	0.013***	0.006
	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.004)
Educational time others	0.017	0.013	0.012	0.007	0.000	0.000
	(0.016)	(0.015)	(0.015)	(0.016)	(0.017)	(0.015)
Care time parents	-0.000	0.000	0.002	0.003	-0.001	-0.000
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Care time others	0.005	0.003	0.002	0.009	0.012*	0.007
	(0.006)	(0.006)	(0.005)	(0.006)	(0.006)	(0.006)
Lagged test outcome		0.456***		0.178***		0.455***
		(0.019)		(0.039)		(0.020)
Observations	2,753	2,690	7,428	2,504	2,399	2,392
Lagged dependent	NO	YES	NO	YES	NO	YES
Lagged inputs	NO	NO	NO	NO	YES	YES
Fixed effects	NO	NO	YES	YES	NO	NO

Notes: Cognitive skills are measured using Matrix Reasoning Test outcomes standardized by age group. Each specification includes controls as the age of the primary caregiver, the number of siblings, the log of family income, and dummies for the college education of the primary caregiver, the presence of both biological parents at home, the gender of the study child, the indigenous status of the study child, the study child's cohort, and the day of data collection.

Table B.48: Estimated parameters of production function with jointly estimated parenting dimensions for cognitive skills (MRT) at age 10-11

	OLS	VA	FE	AB	CU	CV
Warm style	-0.018	-0.008	-0.023	-0.021	-0.012	-0.007
	(0.013)	(0.011)	(0.018)	(0.027)	(0.031)	(0.025)
Reasoning style	-0.062***	-0.035***	-0.024	0.008	-0.047*	-0.034
	(0.013)	(0.011)	(0.018)	(0.028)	(0.027)	(0.023)
Hostile/inconstistent style	-0.040***	-0.021*	0.001	0.013	-0.005	0.000
	(0.013)	(0.011)	(0.018)	(0.025)	(0.028)	(0.023)
Consistent style	0.007	-0.001	0.011	0.024	0.048*	0.031
	(0.013)	(0.011)	(0.018)	(0.024)	(0.025)	(0.022)
Educational time parents	-0.001	-0.002	-0.002	0.002	0.001	0.002
	(0.003)	(0.002)	(0.004)	(0.004)	(0.005)	(0.004)
Educational time others	0.003	0.002	-0.009	-0.001	0.006	-0.001
	(0.005)	(0.004)	(0.008)	(0.012)	(0.008)	(0.005)
Care time parents	0.001	0.001	0.004**	0.003	0.001	0.002
	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Care time others	0.005**	0.005***	0.007**	0.003	0.003	0.004
	(0.002)	(0.002)	(0.003)	(0.004)	(0.004)	(0.003)
Lagged test outcome		0.496***		0.178***		0.463***
		(0.012)		(0.039)		(0.019)
Observations	7,266	7,055	7,428	2,504	2,262	2,256
Lagged dependent	NO	YES	NO	YES	NO	YES
Lagged inputs	NO	NO	NO	NO	YES	YES
Fixed effects	NO	NO	YES	YES	NO	NO

Notes: Cognitive skills are measured using Matrix Reasoning Test outcomes standardized by age group. Each specification includes controls as the age of the primary caregiver, the number of siblings, the log of family income, and dummies for the college education of the primary caregiver, the presence of both biological parents at home, the gender of the study child, the indigenous status of the study child, the study child's cohort, and the day of data collection.

Table B.49: Estimated parameters of production function with jointly estimated parenting dimensions for cognitive skills (PPVT) at age 8-9

	OLS	VA	FE	AB	CU	CV
Warm style	-0.034*	-0.033**	-0.052***	-0.038*	-0.032	-0.036
	(0.019)	(0.017)	(0.020)	(0.022)	(0.026)	(0.022)
Reasoning style	-0.084***	-0.025	0.014	0.007	-0.040	-0.018
	(0.021)	(0.018)	(0.021)	(0.022)	(0.028)	(0.024)
Hostile/inconstistent style	-0.032*	-0.025	-0.026	-0.015	-0.004	-0.005
	(0.019)	(0.016)	(0.021)	(0.023)	(0.025)	(0.021)
Consistent style	-0.002	0.005	0.025	0.018	0.014	0.030
	(0.019)	(0.017)	(0.021)	(0.022)	(0.023)	(0.020)
Educational time parents	0.016***	0.010***	0.001	-0.004	0.015***	0.008**
	(0.004)	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)
Educational time others	0.013	0.001	-0.016	0.002	0.002	-0.008
	(0.015)	(0.015)	(0.025)	(0.025)	(0.016)	(0.016)
Care time parents	0.002	0.000	-0.001	-0.002	-0.001	-0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Care time others	-0.001	-0.003	-0.004	0.003	0.006	0.001
	(0.005)	(0.005)	(0.010)	(0.011)	(0.006)	(0.006)
Lagged test outcome		0.494***		0.151***		0.490***
		(0.019)		(0.044)		(0.020)
Observations	2,755	2,633	3,437	2,156	2,401	2,343
Lagged dependent	NO	YES	NO	YES	NO	YES
Lagged inputs	NO	NO	NO	NO	YES	YES
Fixed effects	NO	NO	YES	YES	NO	NO

Notes: Cognitive skills are measured using Peabody Picture Vocabulary Test (PPVT) outcomes standardized by age group. Each specification includes controls as the age of the primary caregiver, the number of siblings, the log of family income, and dummies for the college education of the primary caregiver, the presence of both biological parents at home, the gender of the study child, the indigenous status of the study child, the study child's cohort, and the day of data collection.

Table B.50: Estimated parameters of production function with extended set of controls for cognitive skills (MRT) at age 8-9

	OLS	VA	FE	AB	CU	CV
Parental warmth	-0.035	-0.028	-0.033	-0.015	-0.030	-0.029
	(0.027)	(0.025)	(0.024)	(0.033)	(0.034)	(0.031)
Parental reasoning	-0.016	-0.015	-0.007	-0.011	-0.027	-0.025
	(0.024)	(0.021)	(0.021)	(0.028)	(0.028)	(0.025)
Hostile parenting	-0.006	0.006	0.013	0.002	-0.005	0.010
	(0.026)	(0.023)	(0.023)	(0.030)	(0.032)	(0.027)
Inconsistent parenting	-0.061**	-0.045**	-0.027	0.002	-0.025	-0.021
	(0.025)	(0.022)	(0.024)	(0.031)	(0.031)	(0.027)
Attempted consistency	0.007	-0.011	-0.032*	-0.029	-0.026	-0.046*
	(0.022)	(0.020)	(0.019)	(0.025)	(0.027)	(0.023)
Educational time parents	0.013***	0.006*	0.000	0.003	0.013***	0.006
	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.004)
Educational time others	0.018	0.015	0.013	0.008	0.007	0.005
	(0.016)	(0.015)	(0.015)	(0.016)	(0.017)	(0.016)
Care time parents	-0.000	0.000	0.002	0.003	-0.001	-0.000
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Care time others	0.005	0.002	0.001	0.009	0.012*	0.005
	(0.006)	(0.006)	(0.005)	(0.006)	(0.007)	(0.007)
Lagged test outcome		0.448***		0.180***		0.447***
		(0.019)		(0.039)		(0.021)
Observations	2,709	2,649	7,428	2,503	2,364	2,357
Lagged dependent	NO	YES	NO	YES	NO	YES
Lagged inputs	NO	NO	NO	NO	YES	YES
Fixed effects	NO	NO	YES	YES	NO	NO

Notes: Cognitive skills are measured using Matrix Reasoning Test outcomes standardized by age group. Each specification includes controls as the age of the primary caregiver, the number of siblings, the log of family income, standardized Social-Economic Indexes for Areas, weight at birth and dummies for the college education of the primary caregiver, the presence of both biological parents at home, the gender of the study child, the indigenous status of th study child, the study child's cohort, the day of data collection, urban area, Australian state and early birth.

Table B.51: Estimated parameters of production function with extended set of controls for cognitive skills (MRT) at age 10-11

	OLS	VA	FE	AB	CU	CV
Parental warmth	-0.031*	-0.014	-0.026	-0.019	-0.022	-0.011
	(0.017)	(0.015)	(0.022)	(0.031)	(0.035)	(0.029)
Parental reasoning	0.002	0.008	-0.006	-0.017	-0.039	-0.030
	(0.015)	(0.013)	(0.021)	(0.028)	(0.028)	(0.025)
Hostile parenting	-0.002	0.007	0.010	0.013	0.013	0.015
	(0.017)	(0.014)	(0.022)	(0.030)	(0.030)	(0.026)
Inconsistent parenting	-0.077***	-0.051***	-0.031	0.005	-0.047	-0.036
	(0.016)	(0.013)	(0.020)	(0.030)	(0.032)	(0.027)
Attempted consistency	0.015	0.004	0.013	-0.002	0.020	0.023
	(0.014)	(0.012)	(0.020)	(0.026)	(0.027)	(0.023)
Educational time parents	0.000	-0.001	-0.002	0.001	0.002	0.003
	(0.003)	(0.002)	(0.004)	(0.005)	(0.005)	(0.004)
Educational time others	0.003	0.002	-0.008	0.001	0.006	-0.001
	(0.005)	(0.004)	(0.009)	(0.012)	(0.008)	(0.006)
Care time parents	0.000	0.000	0.004**	0.003	0.000	0.001
	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Care time others	0.004**	0.005***	0.006**	0.003	0.001	0.002
	(0.002)	(0.002)	(0.003)	(0.004)	(0.004)	(0.003)
Lagged test outcome		0.491***		0.180***		0.456***
		(0.012)		(0.039)		(0.020)
Observations	7,168	6,965	7,428	2,503	2,224	2,218
Lagged dependent	NO	YES	NO	YES	NO	YES
Lagged inputs	NO	NO	NO	NO	YES	YES
Fixed effects	NO	NO	YES	YES	NO	NO

Notes: Cognitive skills are measured using Matrix Reasoning Test outcomes standardized by age group. Each specification includes controls as the age of the primary caregiver, the number of siblings, the log of family income, standardized Social-Economic Indexes for Areas, weight at birth and dummies for the college education of the primary caregiver, the presence of both biological parents at home, the gender of the study child, the indigenous status of th study child, the study child's cohort, the day of data collection, urban area, Australian state and early birth.

Table B.52: Estimated parameters of production function with extended set of controls for cognitive skills (PPVT) at age 8-9

	OLS	VA	FE	AB	CU	CV
Parental warmth	-0.049**	-0.045**	-0.069***	-0.055*	-0.041	-0.041
	(0.024)	(0.021)	(0.026)	(0.028)	(0.031)	(0.026)
Parental reasoning	0.016	0.011	-0.002	-0.003	-0.006	-0.019
	(0.023)	(0.020)	(0.025)	(0.026)	(0.028)	(0.023)
Hostile parenting	-0.014	-0.022	-0.041	-0.035	-0.000	-0.011
	(0.023)	(0.021)	(0.025)	(0.027)	(0.028)	(0.025)
Inconsistent parenting	-0.079***	-0.029	0.014	0.020	-0.038	-0.019
	(0.023)	(0.021)	(0.026)	(0.026)	(0.029)	(0.026)
Attempted consistency	0.027	-0.002	-0.011	0.003	0.004	0.002
	(0.020)	(0.018)	(0.021)	(0.023)	(0.024)	(0.021)
Educational time parents	0.015***	0.009**	0.001	-0.003	0.013***	0.007*
	(0.004)	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)
Educational time others	0.014	0.000	-0.015	0.002	0.005	-0.005
	(0.014)	(0.015)	(0.025)	(0.025)	(0.016)	(0.016)
Care time parents	0.002	0.000	-0.001	-0.002	-0.001	-0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Care time others	0.000	-0.002	-0.005	0.003	0.007	0.001
	(0.005)	(0.006)	(0.010)	(0.011)	(0.006)	(0.006)
Lagged test outcome		0.481***		0.150***		0.478***
		(0.019)		(0.043)		(0.021)
Observations	2,711	2,593	3,437	2,156	2,366	2,308
Lagged dependent	NO	YES	NO	YES	NO	YES
Lagged inputs	NO	NO	NO	NO	YES	YES
Fixed effects	NO	NO	YES	YES	NO	NO

Notes: Cognitive skills are measured using Peabody Picture Vocabulary Test (PPVT) outcomes standardized by age group. Each specification includes controls as the age of the primary caregiver, the number of siblings, the log of family income, standardized Social-Economic Indexes for Areas, weight at birth and dummies for the college education of the primary caregiver, the presence of both biological parents at home, the gender of the study child, the indigenous status of th study child, the study child's cohort, the day of data collection, urban area, Australian state and early birth.

C. APPENDIX TO CHAPTER 3

C.1. Additional tables

Table C.1: Logistic Regression for Marriage Migration

	Better	Worse
Log of sex ratio	3.937***	4.599***
	(0.494)	(0.337)
Individual characteristics:		
Age	0.001	0.009
	(0.006)	(0.007)
Primary educ. or higher	0.191***	0.323***
	(0.065)	(0.068)
Household characteristics:		
Head with primary educ. or higher	0.050	0.256***
	(0.064)	(0.066)
Urban	0.085	-0.197**
	(0.073)	(0.084)
Scheduled Tribe	-0.701***	-0.256*
	(0.115)	(0.136)
Scheduled Caste	-0.010	0.139
	(0.086)	(0.088)
Other backward classes	0.096	0.219***
	(0.068)	(0.072)
Log of consumption per capita	0.005	0.018
	(0.070)	(0.069)
Landowners	-0.030	0.035
	(0.061)	(0.068)
Observations	17,138	15,270
State FE	No	No

Notes: 1. The sample consists of all women aged 20-35 who did not move out of their birth state. 2. Sex ratio is the number of men per 100 women in the marriageable group in the current district. 3. Household consumption is defined as monthly household expenditures per capita in thousands of Rupees. 4. Household owns land if the acreage possessed is more than 0.01 hectares. *Source:* Authors' calculations from the NSS 2006/7. $*p < 0.1, **p < 0.05, ***p < 0.01$.

Table C.2: Logistic Regression for Marriage Migration

	Rural	Urban
Log of sex ratio	4.220***	2.411***
	(0.222)	(0.320)
Individual characteristics:		
Age	0.000	-0.006
	(0.004)	(0.007)
Primary educ. or higher	0.205***	0.344***
	(0.043)	(0.079)
Household characteristics:		
Head with primary educ. or higher	0.146***	0.195**
	(0.041)	(0.080)
Scheduled Tribe	-0.326***	-0.225
	(0.076)	(0.170)
Scheduled Caste	0.003	0.116
	(0.057)	(0.094)
Other backward classes	0.019	0.235***
	(0.047)	(0.070)
Log of consumption per capita	0.135***	-0.086
	(0.049)	(0.064)
Landowners	-0.037	0.145**
	(0.040)	(0.070)
Observations	35,347	16,428
State FE	No	No

Notes: 1. The sample consists of all women aged 20-35 who did not move out of their birth state. 2. Sex ratio is the number of men per 100 women in the marriageable group in the current district. 3. Household consumption is defined as monthly household expenditures per capita in thousands of Rupees. 4. Household owns land if the acreage possessed is more than 0.01 hectares. *Source:* Authors' calculations from the NSS 2006/7. $*p < 0.1, **p < 0.05, ***p < 0.01$.