Parental Networks, Wage Expectations, and
the Intergenerational Educational Mobility[*]

Erdenebulgan Damdinsuren[†], Mariya Mitkova[‡], Anna Zaharieva [§]

August 11, 2022

Abstract

We develop a theoretical labour market model with two generations of workers, endogenous social networks of parents and binary schooling choices of children. Since the market skill premium is unobservable, families rely on noisy wage information obtained from their social contacts giving rise to heterogeneous expectations across families. If social networks are subject to skill homophily and high skill parents are a minority, then children in low skill families are stronger affected by the lack of objective information, their expectations are more dispersed and they are less likely to study giving rise to a positive intergenerational schooling correlation. Next, we calibrate the model to the German labour market data (SOEP) and show that the described mechanism can potentially account for up to 15% of the intergenerational schooling correlation. Moreover, the data strongly supports the idea of skill homophily in the social networks. Finally, we test the main microeconomic prediction of the model that a larger share of high skill parental friends is positively associated with a probability of obtaining higher education (with a declining marginal gain). We find that this prediction is inline with the empirical data (SOEP) even after controlling for the education of parents, cognitive abilities of parents and children, demographics and some non-cognitive characteristics (locus of control and personality traits).

Keywords: intergenerational mobility, social networks, heterogeneous expectations

JEL Classification: D10, I20, J30, J62

[†]E-mail: edamdinsuren@uni-bielefeld.de Tel.: +49-521-106-4864, Bielefeld Graduate School of Economics and Business, Bielefeld University, 33615 Bielefeld, Germany; Department of Economics and Finance, Catholic University of the Sacred Heart, 20123 Milan, Italy.

[‡]E-mail: mariya.mitkova@uni-bielefeld.de Tel.: +49-521-106-3942, Department of Business Administration and Economics, Bielefeld University, 33501 Bielefeld, Germany,

[§]E-mail: anna.zaharieva@uni-bielefeld.de Tel.: +49-521-106-5637, Fax: +49-521-106-89005, Center for Mathematical Economics and Department of Business Administration and Economics, Bielefeld University, 33501 Bielefeld, Germany

1 Introduction

Vast empirical research shows that educational outcomes of parents and children are strongly correlated[1]. In the meantime there is consensus in the literature that the intergenerational transmission of schooling is accounted for by a collection of factors and mechanisms. These include genetic transmission of cognitive and non-cognitive skills and abilities, wealth inheritance and financial constraints as well as public policies in the educational system. Nevertheless, Bowles and Gintis (2002, p 4-5.) conclude that the "transmission of economic success across generations remains something of a black box" since existing mechanisms explain "at most three-fifths of the intergenerational transmission of economic status". Thus, in this paper we direct our attention to a novel mechanism and investigate the role of parental social networks for the expectations and educational attainment of children. Our analysis includes a calibrated theoretical model of schooling decisions which we use to guide empirical research based on the German labour force survey data (SOEP).

More specifically, we develop a labour market model with endogenous social networks of parents. In addition, we incorporate recent evidence summarized in French and Oreopoulos (2017) that there is a high degree of uncertainty among the youth concerning the gains and costs of higher education and even the application process itself. So we think of the real return to schooling/skill premium as unknown to the market participants[2]. Hence, young individuals in the model use all available information to obtain an estimate of the return to schooling and use this estimate to make a decision about acquiring higher education or not. More specifically, we assume that the offspring generation uses noisy wage observations of social contacts in parental networks to form their expectations about the market skill premium. In this setting we show that differences in the composition of social networks across families give rise to heterogeneous expectations of children and lead to different educational decisions and labour market outcomes.

When modeling the networks of parents we take into account a common empirical fact that social network formation is subject to skill homophily, meaning that individuals coming from the same skill group are more likely to form social links (see Montgomery (1991) and McPherson et al. (2001)). In a labour market with a binary skill structure this leads to a situation when high skill friends are overrepresented in the social networks of high skill parents, whereas friendship ties with low skill workers are less common in high skill families. The opposite is true for low skill parents. However, the situation of children in the two types of families is not symmetric since high skill workers are a minority and high skill wages tend to be more dispersed than low skill wages. In this setting we show that social networks of high skill parents are more balanced between the two skill groups, thus they possess more precise information about the market skill premium. Intuitively, this means that expectations about the skill premium are less dispersed across high skill families, and so children in these families are more likely to acquire higher education. This

---

[1] We review this literature in the next subsection

[2] Some justification for this assumption can be found in Card (2001), who presents a summary of 11 studies estimating the return to schooling. The table shows a large variation in the estimates depending on the time period, sample restrictions and the estimation techniques. This evidence suggests that a precise estimate of the market skill premium is difficult to obtain. In our model, adolescents collect information about the skill premium from their parents and parental friends, and they put some weights on these sources of information depending on their reliability and precision. Thus, they combine different sources of information with different precision into one estimate, which is used for the final qualification decision. It would be a straightforward extension of the model to accommodate other point estimates obtained, for example, from the internet or mass media.

mechanism gives rise to a positive correlation in the schooling outcomes of parents and children in isolation from other traditional explanations of intergenerational mobility[3].

Concerning expectations, we start with a model where wages of parental social contacts are perfectly observed by the offspring in order to provide the intuition, but we calibrate a modified model where wage observations in the social networks of parents are subject to an observation error. The intuitive idea behind this setting is that children obtain an imprecise signal of the actual labour income of a family friend by observing this friend's living expenses (e.g. large or small house, luxus or simple furniture, expensive or cheap car etc.). This modification creates a situation when some wages are observed precisely (e.g. wages of parents), whereas other wage observations are subject to noise leading to a problem of heteroscedasticity in the expectation formation. We show that this problem can be dealt with by applying a weighted least squares approach implying that children should optimally assign a higher weight to the wage of their parent and lower weights to the remaining wage observations stemming from the network of the parent. This approach gives rise to the efficient estimates of the return to schooling. Overall, however, the intergenerational correlation coefficient is only moderately sensitive to the noise parameter since social networks of parents are sufficiently large and the noise variable is unbiased.

In the next step the modified model is calibrated to match empirical observations in the German labour market based on the Socio-Economic Panel (SOEP). The advantage of using this data is that it includes information on social networks combined with education and labour market outcomes of parents and children. We find that the German empirical data on social networks strongly supports the idea of skill homophily. For example, the share of high skill friends in the networks of high skill workers is almost 20% above the population average, whereas the share of high skill friends in the networks of low skill workers is almost 20% below the population average. Based on this information we estimate a model-implied skill homophily paramter and find that it is remarkably stable over time. The data also supports the assumptions that high skill workers with an upper secondary school degree certificate ("Abitur") are a minority in the labour market (24.8% in the generation of parents and 37.8% in the generation of children) and their wages are more dispersed compared to low skill workers. The latter observation has an amplifying effect on the intergenerational transmission of schooling in our model. Based on the performed calibration we find that the described mechanism building on parental social networks and labour market uncertainty can potentially explain up to $15\% = (0.055/0.375)$ of the observed schooling correlation.

Further, we use the calibration to analyze the influence of the key model parameters on intergenerational mobility and to evaluate the sensitivity of different mobility measures to these parameters. We find that the cost of education, network size, skill homophily and the share of high skill families have an inverse U-shape impact on the intergenerational correlation coefficient and the mobility ratio. Quantitatively, however, the effect of the network size is small, whereas both mobility measures are sensitive to the cost of education, the skill homophily and the skill

---

[3]In the appendix we present an extension of the model allowing for the transmission of abilities, which can be genetically inherited (nature) or transmitted in early childhood (nurture). The purpose of this extension is to analyze the interaction of the new network-driven and the traditional ability-driven components in shaping the intergenerational schooling correlation. In this extension we prove theoretically that the interaction effect is always negative. As the ability correlation of parents and children gets stronger, the number of high ability children born into low skill families is declining making the informational disadvantage of low skill families less important and crowing out the network effect.

composition of parental networks. Given that the cost of education (monetary expenses and effort) is generally unobservable and the homophily parameter is found to be stable over the last two decades, we focus on the skill composition of parental friends to test the model prediction.

Our empirical data (SOEP) includes rare information about the education of parents, children and parental friends, limited to the three best friends of each parent. Based on this data we construct a measure corresponding to the share of high skill friends in the networks of parents and include it into the logistic regression predicting a binary educational outcome of the child. Thus, we follow the econometric approach summarized in Plug et al. (2018) and use the education of parents as well as some demographic characteristics of the child as control variables. The estimated coefficient corresponding to the share of high skill family friends has a causal interpretation only if the model does not suffer from an omitted variable problem meaning that the set of regressors should include variables related to the formation of parental networks. Thus, we additionally control for cognitive abilities of parents and children as well as some non-cognitive characteristics including the locus of control and the big five personality traits (see McCrae and Costa (1999) and Anger and Heineck (2010)). We find that the coefficient on the share of high skill family friends is positive and significant but the marginal gain associated with this effect is declining. This indicates that children exposed to interactions with a higher number of educated family friends are more likely to obtain higher education. We take this result as first evidence supporting the theoretical model, yet, we note that there are multiple theoretical mechanisms that can be compatible with the observed evidence. Thus, future research in this direction should focus on comparing these mechanisms and evaluating their relative contributions to the positive link between the quality of parental networks and the educational outcomes of the youth.

1.1   Related literature

Our study is related to several strands of literature. Empirical work on measuring intergenerational correlations between the income/schooling levels of parents and their children in a multi-country framework includes Björklund and Jäntti (1997), Chevalier et al. (2003), Hertz et al. (2007), Bratberg et al (2016) and Leone (2019). Hertz et al. (2007) provide an extensive survey of estimates on the parent-child educational correlations. Their estimates vary from 0.45-0.55 for low mobility countries (Italy, Slovenia, Austria, Ireland, USA, and Switzerland) down to 0.3-0.4 for countries with high educational mobility (Sweden, Belgium, Netherlands, Norway, Finland, UK, and Denmark). Björklund and Jäntti (1997) and Bratberg et al. (2016) confirm the finding that intergenerational mobility is lower in the USA compared to Sweden and Norway. Empirical evidence for Germany is provided in Couch and Dunn (1997), Heineck and Riphahn (2009), Riphahn and Schieferdecker (2012) and Leone (2019). The former study utilizes early waves of the Socio-economic Panel and reports the estimates of the intergenerational schooling correlations in the range $0.237 - 0.39$ depending on the gender of the parent and the child, which puts Germany into a group of studies with relatively high intergenerational mobility. Empirical estimates reported in Leone (2019) support this view.

The literature proposes several explanations for the observed correlations (see Piketty (2000) for more details). First, cognitive and non-cognitive skills and abilities, as well as preferences and social norms, may be inherited by children from their parents (Becker and Tomes (1986)). On

the one hand, abilities can be transmitted genetically from parents to their biological children (nature). On the other hand, better educated parents are more efficient in transmitting knowledge and supporting their children in the studying process (nurture). Several empirical studies report strong correlations in cognitive abilities of parents and children based on IQ test scores highlighting the importance of a combined effect of nature and nurture (Black et al. (2009), Björklund et al. (2010)). Sacerdote (2007) undertakes an attempt to separate the contribution of nature and nurture by comparing educational outcomes of adoptive and biological children and reports that the genetic transmission of abilities contributes more to explaining educational attainment of children. The study by Holmund et al. (2011) comes to a similar conclusion that parental nurture effect does not play a large role.

Second, if the credit market is not perfect, financial constraints may prevent children in poor families from acquiring education (Galor and Zeira (1993)). Yet, empirical support for this transmission channel is rather weak, for example, Carneiro and Heckman (2002) and Belley et al. (2011) find little evidence that short term credit constraints explain a part of the gap in college enrollment between children from high and low income families. Third, the institutional design of the education system may play a crucial role in the attainment of education (Checchi et al; 1999, Checchi and Flabbi; 2007). These studies argue that countries with education systems characterized by equal access to education are more likely to have higher educational mobility, and the public expenditure could be one of the effective factors for increasing equal opportunity in education. For example, Chevalier et al. (2003) find a negative relationship between the intergenerational schooling correlation and the share of public expenditure in tertiary education.

Forth, and most closely related to our research is the explanation that children's education levels and their labor market outcomes can also be affected by the neighborhood or social environment where they grow up (Chetty et al. (2014, 2016), Katz et al. (2001)). Here one can distinguish between the social and the geographical dimension even though the two are closely related. The geographical component of a living environment (e.g. quality of schools, family safety, access to important infrastructure and facilities) is intensively studied in the literature on neighborhood effects. These studies include Patacchini and Zenou (2011), Del Bello et al. (2015), Chetty et al. (2016), Altonji and Mansfield (2018), and Lavecchia et al. (2020). On the one hand, Patacchini and Zenou (2011) find that parental involvement effect is stronger for educational attainment by children in high skill families. While the neighborhood effect is more important for children in low skill families. Also Altonji and Mansfield (2018) and Lavecchia et al (2020) suggest that the quality of school and neighborhood could lead to an increase in the share of graduating students. But on the other hand, Chetty et al. (2016) analyze the impact of the "Moving to Opportunity" experiment in the USA and find that moving to a better neighborhood has negative impact on the educational outcomes of adolescences and a moderately positive impact for children who were young at the time of the family move.

The literature on the network dimension of the living environment is rather scarce. Several empirical studies investigate the effect of parental job referrals on their children's labour market outcomes. For example, Kramarz and Skans (2014) analyze Swedish data and find that the parental networks effects are stronger for lower educated children when they seek their first jobs. Corak and Piraino (2011) consider the intergenerational transmission of employers and document that the probability of having the same employer for parents and children is higher

for the top-earning groups in Canadian data. Going beyond referral hiring Plug et al. (2018) investigate the effect of parental friendships on the labor market outcomes of children. They find a small positive effect on the occupational choices of children at the time of labour market entry and an insignificant effect at later stages of the child's career. Comparing the network and the geographic components of the living environment Del Bello et al. (2015) suggest that the peer effect on education is stronger than the effect of neighborhood.

The study which is most closely related to our theoretical research is by Calvó-Armengol and Jackson (2009). The key element of their model is an assumed complementarity in the actions of players, meaning that a person is more likely to choose a high action (e.g. higher education) if there are many actors in the surrounding environment taking the same action. This assumption leads to situations when the actions of parents and children are correlated due to the common exposure to the same social environment without a direct transmission of skills or abilities from parents to children. In this paper we continue this strand of research but deviate from the exogenous complementarity assumption. Instead, we look deeper at the underlying mechanisms of network and expectation formation in a situation when the market skill premium is unobservable giving rise to heterogeneous expectations and correlated schooling outcomes of parents and children. Hence we contribute to the described literature in the following ways: 1) by advancing and extending the theoretical framework modeling the link between social networks and intergenerational mobility 2) by analyzing the impact of network structure (e.g. network size and skill homophily) and labour market uncertainty on intergenerational mobility measures, which is only possible if the network structure is modeled explicitly 3) by quantifying the potential contribution of the developed mechanism based on the calibrated model 4) by examining the impact of parental networks on the educational attainment of children based on the household survey data.

Finally, our paper is broadly related to the large theoretical literature on the role of social networks and families for labour market outcomes. These studies include but are not limited to Calvó-Armengol and Zenou (2005), Ioannides and Soetevent (2006), Fontaine (2008), Cahuc and Fontaine (2009), Galenianos (2013), Galeotti and Merlino (2014) as well as Stupnytska and Zaharieva (2015, 2017). The primary focus of these papers is on the impact of referral hiring among connected workers. Seminal study introducing the idea of network homophily into economic modeling is Montgomery (1991). Several extensions of this approach focusing on the role of network homophily are conducted in Horvath (2014) and Zaharieva (2018). The role of family contacts in a theoretical framework is analyzed by Ek and Holmlund (2010), Flabbi and Mabli (2018) and Zaharieva (2013). In addition, Doepke and Tertilt (2016) analyze the impact of families in a broader macroeconomic context and show that accounting for the family leads to new answers to classical economic questions. Even though these papers make multiple contributions and extend our knowledge about families and social networks, none of them addresses the implications of networks for intergenerational mobility.

The article is organized as follows. Section 2 presents the core theoretical model with endogenous social networks, heterogeneous expectations of the offspring generation and their educational choices. Section 3 describes the data and the calibration approach. Section 4 presents a number of numerical results complementing the theory. Section 5 contains empirical analysis. Section 6 concludes the paper.

2 The Model

In this section we develop the analytical model and consider the choice of schooling from an intergenerational perspective. There are two generations of workers: parents (first generation) and children (second generation). We abstract from the fertility issues and assume "one parent - one child"' family structure. In section 2.1 we explain how the social networks of parents are formed. Further, in section 2.2 we show how the generation of children forms expectations about the market return to schooling (skill premium) by observing wages in the networks of their parents. Based on the estimated return to schooling, inborn ability and the corresponding cost parameter they decide to become high or low skilled, thus we model schooling as a binary choice variable. Parents and children share the same information and their goals are aligned.

First generation workers are heterogeneous with respect to their skill, hence some children are born in low skill families L, while others are born in high skill families H. Let $h < 0.5$ denote the fraction of high skill families/parents and $1 - h$ – fraction of low skill families/parents. All parents are employed in the labour market and receive (log)wages w_i^L or w_i^H depending on their skill level[4]. Wage differences among workers with identical skills could stem from the hetereogeneity of employers or reflect different wage strategies of identical firms (e.g. endogenous wage dispersion in the spirit of Burdett and Mortensen (1998)). We interpret these (log)wages as life-time income and assume that they are normally distributed with the corresponding means and variances $E[w_i^L]$, $E[w_i^H]$, $V[w_i^L]$, $V[w_i^H]$. High skill workers earn more on average, so that $E[w_i^L] < E[w_i^H]$. The assumption that (log) wages are normally distributed is not restrictive since the model is based on average wages, which would be asymptotically normally distributed under the Central Limit Theorem (given independent draws). However, we also consider a weighted average wage, where the CLT is not applicable.

The inborn ability of children can be high with probability p and low with probability $1 - p$. In the main specification of the model we consider the setup where p is identical in high and low skill families, therefore any correlation in the educational attainment of parents and children is due to differences in the available information about the labour market obtained via the parental network of social contacts. In the appendix we consider an extension of the model where $p^H > p^L$ giving rise to a positive correlation in abilities between parents and children. For this case we show that the intergenerational schooling covariance can be decomposed into the ability component, the network component and the negative interaction term.

For high ability children the cost of obtaining education is denoted by c[5]. We assume that the cost of education is below the skill premium, that is: $c < E[w_i^H] - E[w_i^L]$, thus education is gainful on average for high ability individuals. The cost parameter c includes monetary costs and effort associated with studying. High ability children observe wages and skills among the contacts of their parents and use this information to form an expectation about the market skill premium. This is a benchmark specification. In section 2.4 we consider a modified specification where wages of social contacts are observed with an error/noise. Based on the estimated skill premium and the cost of obtaining education c high ability children decide whether they obtain higher education and become high-skilled workers or not. The cost of obtaining education for

---

[4]In the extended model we allowed for a higher risk of unemployment faced by low skill parents, however, it has negligible impact on the quantitative results and is disregarded in the following for the ease of exposition

[5]The cost is expressed as a % of the average low skill wage to be comparable with the return to schooling

low ability children is prohibitively high, so they do not consider acquiring education.

2.1 Social networks

In this section we analyze network formation in the first generation of workers. In doing so we follow the random matching approach developed in Neugart and Zaharieva (2018). Formation of social links is subject to skill homophily, that is workers are more likely to create social links with others of the same skill type. In general, homophily refers to the fact that people are more likely to maintain relationships with other people who are similar to themselves. There can be homophily measured by age, race, gender, religion or skill and it is generally a robust observation in social networks (see McPherson et al. (2001) for an overview of research on homophily). The focus of this paper is on the latter type of homophily by the level of education/skill. To the best of our knowledge, Montgomery (1991) was the first study that introduced network homophily (the "inbreeding bias") into an analytical labour market model.

At rate ϕ every worker can be randomly matched with another worker. Let τ denote the probability of creating a social link with a worker of the same skill type and $(1 - \alpha)\tau$, $0 < \alpha < 1$ – be the probability of creating a link with a worker of a different skill type (conditional on matching). Note that the special case when $\alpha = 0$ corresponds to the situation without homophily, while $\alpha = 1$ corresponds to the complete network segregation by skill. Hence α corresponds to the degree of skill homophily in the society. In order to keep the model tractable we consider directed links. This means that, if two workers i and j are randomly matched, it is possible that j becomes a social contact of i but not necessarily vice versa. Some justification for this assumption can be found in Plug et al. (2018) who show that less than a half of parental friendship ties are reciprocated. Every social link can be destroyed at rate δ.

Let ε_k^{ij} denote a fraction of type i workers with exactly k social contacts of type j, $i,j \in \{L, H\}$. This is a fraction out of all type i workers. Consider some type L worker without contacts of the same type. With our notation this worker belongs to the group ε_0^{LL}. At rate ϕ this worker is matched with some other individual. With probability $1 - h$ this individual is of the same type L, so the social link is created with probability τ. Next, consider another type L worker with only one contact of the same type belonging to the group ε_1^{LL}. This person may lose the only contact at rate δ. In the steady state, the propensity for the worker to make transitions between the two states $k - 1$ and k will be equalized, this means:

$$
\begin{aligned}
\varepsilon_0^{LL}\phi(1-h)\tau &= \delta\varepsilon_1^{LL} \Rightarrow \varepsilon_1^{LL} = \frac{\varepsilon_0^{LL}\phi(1-h)\tau}{\delta} \\
\varepsilon_1^{LL}\phi(1-h)\tau &= 2\delta\varepsilon_2^{LL} \Rightarrow \varepsilon_2^{LL} = \varepsilon_0^{LL}\left(\frac{\phi(1-h)\tau}{\delta}\right)^2 \frac{1}{2} \\
\varepsilon_{k-1}^{LL}\phi(1-h)\tau &= k\delta\varepsilon_k^{LL} \Rightarrow \varepsilon_k^{LL} = \varepsilon_0^{LL}\left(\frac{\phi(1-h)\tau}{\delta}\right)^k \frac{1}{k!}
\end{aligned}
$$

Let $\varphi \equiv \phi\tau/\delta$ to simplify the notation. Since all fractions ε_k^{LL} should add up to 1 for $k = 0...\infty$ we get: $\varepsilon_0^{LL} = e^{-\varphi(1-h)}$ and the number of type L contacts has a Poisson distribution with parameter $\varphi(1-h)$. This also means that, $\varphi(1-h)$ is the average number of low-skilled contacts in the social network of a randomly chosen low-skilled worker, we denote it by n^{LL}, so that

$n^{LL} = \varphi(1-h)$ and:

$$\varepsilon_k^{LL} = e^{-n^{LL}}(n^{LL})^k \frac{1}{k!}$$

Alternatively, the type L person can be matched with another person of type H, which happens at rate ϕh. Conditional on matching the social link is formed at rate $(1-\alpha)\tau$. Repeating the same steps as above we get:

$$\varepsilon_k^{LH} = e^{-n^{LH}}(n^{LH})^k \frac{1}{k!} \quad \text{where} \quad n^{LH} \equiv \varphi h(1-\alpha)$$

Here n^{LH} is the average number of high-skilled contacts in the social network of a randomly chosen low-skilled worker. Let n^L denote the average network size for type L workers and γ_L be the fraction of type L contacts in their network, so we get:

$$n^L = n^{LL} + n^{LH} = \varphi[(1-h) + h(1-\alpha)] \quad \gamma_L = \frac{(1-h)}{(1-h) + h(1-\alpha)}$$

Using the same approach for type H workers we get:

$$n^H = n^{HH} + n^{HL} = \varphi[h + (1-h)(1-\alpha)] \quad \gamma_H = \frac{h}{h + (1-h)(1-\alpha)}$$

where n^H is the total network size of high-skilled workers and γ_H is a fraction of type H contacts in their network. One can see that the case of full homophily ($\alpha = 1$) leads to the complete segregation of social networks between the two skill groups, that is $\gamma_L = \gamma_H = 1$. In the opposite case without homophily ($\alpha = 0$), the fraction of contacts of the same type is equal to the fraction of this type in the total population, that is $\gamma_L = 1 - h$ and $\gamma_H = h$. We can also see that high-skilled contacts are underrepresented in the networks of low-skilled workers ($1-\gamma_L < h$), whereas they are overrepresented in the networks of high-skilled workers ($\gamma_H > h$):

$$1 - \gamma_L = \frac{h(1-\alpha)}{(1-h) + h(1-\alpha)} = \frac{h}{(1-h)\frac{1}{(1-\alpha)} + h} < h$$

$$\gamma_H = \frac{h}{h + (1-h)(1-\alpha)} = \frac{h}{(1-h)(1-\alpha) + h} > h$$

Recall that h is the fraction of high-skilled workers (parents) in the labour market, so we use it as a point of reference in the above expressions. Comparing the average sizes of social networks for low and high skilled workers we get the following:

Lemma 1: *The networks of high skill workers/parents are more balanced, meaning $n^{LL} - n^{LH} > n^{HL} - n^{HH}$ but they are smaller on average, $n^H < n^L$ compared to the networks of low skill workers/parents.*

Proof:

$$n^L - n^H = \varphi(1-2h)\alpha > 0 \quad \text{for} \quad h < 0.5$$
$$(n^{LL} - n^{LH}) - (n^{HL} - n^{HH}) = \alpha > 0$$

Lemma 1 shows that the networks of low skill families are more extreme compared to the

9

networks of high skill families. There are two reasons for this finding. First, the number of high skill contacts in the networks of low skill families $n^{LH} = \varphi h(1 - \alpha)$ is small because high skill workers are the minority $h < 0.5$. Second, it is small because of skill homophily $0 < \alpha$. At the same time, the number of low skill contacts in the networks of low skill families is large $n^{LL} = \varphi(1 - h)$ because low skill workers are the majority in the labour market.

Finally, we can calculate the average size of the social network in the economy, which is denoted by n:

$$n = (1 - h)n^L + hn^H = \varphi[1 - 2\alpha h(1 - h)]$$

We can see that parameter φ can be interpreted as a maximum average network size in the absence of homophily ($\alpha = 0$). The reason is that any positive homophily $\alpha > 0$ reduces the network size because fewer matches become social ties.

This subsection shows that high and low skill parents have different composition of social networks. Therefore, information about the market skill premium obtained by children through the social network of their parents is likely to vary across families, giving rise to heterogeneous expectations. We continue by analyzing this point in the next section.

2.2 Labour market

In this section we use the network composition of parents derived above and conditional on their skill type and proceed by analyzing the educational decisions of children. In particular, we assume that children observe the skill types and wages of the social contacts of their parents and form expectations about the market skill premium (return to schooling). This is a benchmark specification. In section 2.4 we consider a modified specification where wages of social contacts are observed with an error/noise.

First, let us consider some family with a low skill parent j. Each low-skill parent has n^{LL} and n^{LH} number of low and high skilled contacts, respectively[6]. Let $w_j^{LH} = \sum_{i=1}^{n^{LH}} w_i^H / n^{LH}$ be the average wage of all high-skilled contacts in family j. When considering an average wage of low skill contacts, the wage of the parent w_j^L is an important source of information from the perspective of the child. Thus, we assign a weight $0 < s < 1$ to w_j^L when calculating the average low skill wage estimated by the child in family j. This means:

$$w_j^{LL} = sw_j^L + (1 - s)\frac{1}{n^{LL}} \sum_{i=1}^{n_{LL}} w_i^L$$

Note that $s = 1/(1 + n^{LL})$ would mean that the wage of the parent is equally important as all other wages of low skill contacts (equal weighting). For $s > 1/(1 + n^{LL})$ children assign a higher weight to the wage of the parent, thus variable s captures the importance of the parental wage for the decision of the child. At the same time the weight coefficient s does not introduce a bias in the expectations of children because w_j^L is drawn from the same distribution as wages of the social contacts, so that the expectation is unbiased $E[w_j^{LL}] = E[w_i^L]$. However, we show later that the efficiency (variance) of w_j^{LL} is influenced by s.

---

[6]These numbers are averages, however, we take them as an approximation for differences in the network composition across families and ignore the stochasticity of the number of social contacts in the following analysis

Considering the population of all families with low skill parents and assuming that wage draws are independent implies that the average high skill wage w_j^{LH} is normally distributed with its mean and variance:

$$w_j^{LH} = \frac{1}{n^{LH}} \sum_{i=1}^{n^{LH}} w_i^H \sim N(E[w_j^{LH}], V[w_j^{LH}])$$

$$\text{with} \quad E[w_j^{LH}] = E[w_i^H] \quad \text{and} \quad V[w_j^{LH}] = \frac{V[w_i^H]}{n^{LH}}$$

In a similar way, the average low skill wage w^{LL} is normally distributed with the following mean and variance:

$$w_j^{LL} = sw_j^L + (1-s)\frac{1}{n_{LL}} \sum_{i=1}^{n_{LL}} w_i^L \sim N(E[w_j^{LL}], V[w_j^{LL}])$$

$$\text{with} \quad E[w_j^{LL}] = E[w_i^L] \quad \text{and} \quad V[w_j^{LL}] = V[w_i^L]\left(s^2 + \frac{(1-s)^2}{n^{LL}}\right)$$

Note that the underlying normality assumption of (log) wages is necessary for $s \neq 0$ and $s \neq 1/(1 + n^{LL})$, however, it becomes redundant for $s = 0$ or $s = 1/(1 + n^{LL})$, when the Central Limit Theorem would apply.

Let $\Delta_j^L = w_j^{LH} - w_j^{LL}$ be a skill premium estimated by family j (with a low skill parent). Its distribution in the population of low skill families can be written as:

$$\Delta_j^L \sim N\left(E[w_i^H] - E[w_i^L], \frac{V[w_i^H]}{n^{LH}} + V[w_i^L]\left(s^2 + \frac{(1-s)^2}{n^{LL}}\right)\right)$$

Let the corresponding cumulative distribution function be denoted by Φ^L. Note the following:

(a) $\quad V[\Delta_j^L] \quad = \frac{V[w_i^H]}{n^{LH}} + \frac{V[w_i^L]}{n^{LL}} \quad$ if $\quad s = 0$

(b) $\quad V[\Delta_j^L] \quad = \frac{V[w_i^H]}{n^{LH}} + \frac{V[w_i^L]}{n^{LL}+1} \quad$ if $\quad s = \dfrac{1}{n^{LL}+1}$

(c) $\quad V[\Delta_j^L] \quad = \frac{V[w_i^H]}{n^{LH}} + V[w_j^L] \quad$ if $\quad s = 1$

This shows that in the extreme case when $s = 1$ second generation individuals take their parent's wage as the only observation and ignore wages of the low skill contacts of their family. In this case the estimates of the skill premium Δ_j^L are likely to be very dispersed.

Next, consider high-skill families. The distribution of the skill premium Δ_j^H estimated by these families is given by:

$$\Delta_j^H \sim N\left(E[w_i^H] - E[w_i^L], V[w_i^H]\left(s^2 + \frac{(1-s)^2}{n^{HH}}\right) + \frac{V[w_i^L]}{n^{HL}}\right)$$

where n^{HH} is the number of high skill contacts and n^{HL} is the number of low skill contacts of the family/parent. Let the corresponding cumulative distribution function be denoted by Φ^H.

Comparing the distributions for Δ_j^L and Δ_j^H we can see that the estimates of the skill premium/return to schooling are unbiased in both types of families, however, the two variances can be different. The reason is that the precision of information available to children in the two different types of families depends on the structure of social networks, that is the network size

11

parameters $n^{HH} = \varphi h$, $n^{HL} = \varphi(1-h)(1-\alpha)$, $n^{LL} = \varphi(1-h)$ and $n^{LH} = \varphi h(1-\alpha)$. In proposition 1 we compare the two variances and summarize our findings:

Proposition 1: *The skill premiums Δ_j^L and Δ_j^H have the same means, i.e. $E[\Delta_j^L] = E[\Delta_j^H] = E[w_i^H] - E[w_i^L]$. Let $n^{LH} = \varphi h(1-\alpha) > 1$, then the variance of the skill premium Δ_j^L is greater than that of Δ_j^H when:*

$$1 < \frac{V[w_i^H]}{V[w_i^L]} \frac{(1-h)}{h} < 1 + \frac{\varphi(1-\alpha)(1-2h)}{\varphi h(1-\alpha) - 1}$$

Proof: Appendix

Proposition 1 shows that the estimated return to schooling is more dispersed in low skill families, implying more uncertainty, if the weighted variance ratio $V[w_i^H](1-h)/(V[w_i^L]h)$ is above 1 but below an endogenous upper bound. In order to understand this condition consider a simplified case when $s = 0$, that is the wage of the parent is not included in the estimation. The two variances then become:

$$V[\Delta_j^L] = \frac{V[w_i^H]}{\varphi h(1-\alpha)} + \frac{V[w_i^L]}{\varphi(1-h)} \qquad V[\Delta_j^H] = \frac{V[w_i^H]}{\varphi h} + \frac{V[w_i^L]}{\varphi(1-h)(1-\alpha)}$$

and their difference can be written as:

$$V[\Delta_j^L] - V[\Delta_j^H] = \frac{\alpha}{\varphi(1-\alpha)} \left(\frac{V[w_i^H]}{h} - \frac{V[w_i^L]}{(1-h)} \right)$$

Thus the skill premium Δ_j^L is more dispersed if the weighted variance ratio $V[w_i^H](1-h)/(V[w_i^L]h)$ is above 1. In the following we assume that it holds until the end of the paper. This condition becomes more likely when wages of high (low) skill workers are more (less) dispersed. Moreover, the gap in variances $V[\Delta_j^L] - V[\Delta_j^H]$ is larger when network homophily is stronger (higher α) and when social networks are smaller (smaller φ). In the appendix we show that this condition is sufficient but not always necessary for small values of s above zero but below a given threshold.

However, when s is large, meaning that children overweight the importance of parental wages, the situation is different. In order to understand this case consider an extreme point $s = 1$. The two variances are then given by:

$$V[\Delta_j^L] = \frac{V[w_i^H]}{\varphi h(1-\alpha)} + V[w_i^L] \qquad V[\Delta_j^H] = V[w_i^H] + \frac{V[w_i^L]}{\varphi(1-h)(1-\alpha)}$$

and their difference can be written as:

$$V[\Delta_j^L] - V[\Delta_j^H] = -\frac{1}{\varphi(1-\alpha)} \left(\frac{V[w_i^H]}{h}(\varphi h(1-\alpha) - 1) - \frac{V[w_i^L]}{1-h}(\varphi(1-h)(1-\alpha) - 1) \right)$$

This condition gives rise to the upper bound of the weighted variance ratio described in proposition 1. Intuitively, if $s = 1$ then children rely exclusively on the wage of their parent when forming expectations about wages in a respective skill group. If high skill wages are much more dispersed than low skill wages violating the upper bound condition, then children in high skill families possess less precise information about the skill premium than children in low skill families, meaning that $V[\Delta_j^H] > V[\Delta_j^L]$. Therefore, the upper bound condition is necessary and

sufficient for $V[\Delta_j^H] < V[\Delta_j^L]$ when $s = 1$. In the appendix we show that this condition is sufficient (but not always necessary) for large values of s above a given threshold and below 1. Thus proposition 1 describes a set of sufficient conditions leading to a situation when the skill premium estimated by the low skill families Δ_j^L is more dispersed and uncertain compared to the skill premium estimated by the high skill families Δ_j^H.

Next we investigate the decision to acquire education. Consider the generation of children in low skill families, they decide to obtain higher education if the cost c is low: $c < \Delta_j^L$, so the number of children obtaining education in low skill families is $p(1 - \Phi^L(c))$. The corresponding number in high skill families is $p(1 - \Phi^H(c))$. Both distributions $\Phi^H(.)$ and $\Phi^L(.)$ are normal with the same means $E[w_i^H] - E[w_i^L]$. Under the conditions described in proposition 1 the distribution $\Phi^L(.)$ is more dispersed than $\Phi^H(.)$, thus we can say that $\Phi^L(.)$ is obtained from $\Phi^H(.)$ by a mean preserving increase in spread. Moreover, the two cumulative distribution functions satisfy a single-crossing property and intersect only once at the mean. Following Diamond and Stiglitz (1973) this property implies that:

$$\Phi^L(x) - \Phi^H(x) \le (\ge)0 \quad \text{when} \quad x \ge (\le)E[w_i^H] - E[w_i^L]$$

Therefore, we know that $\Phi^L(c) > \Phi^H(c)$ since the cost of acquiring education c is smaller than the wage premium: $c < E[w_i^H] - E[w_i^L]$. This shows the main result of our study: (high ability) children in low skill families face stronger uncertainty when forming expectations about the skill premium, so they are less likely to obtain higher education compared to (high ability) children in high skill families who possess more precise information: $p(1 - \Phi^L(c)) < p(1 - \Phi^H(c))$. The intuition for this result can be explained in the following way: low skill workers are the majority in the market, thus all families have many contacts of this type and the estimate of the average low skill wage is relatively precise. The major difference comes in the estimate of the average high skill wage. High skill workers are the minority, moreover, there are relatively few of them in the networks of low skill families due to the skill homophily, thus the average high skill wage is estimated with low precision and high degree of uncertainty in low skill families. The high degree of uncertainty in turn implies that high ability children in low skill families are more likely to make a mistake by not obtaining the education compared to children in high skill families supplied with more accurate information from the high skill contacts of their parents. This mechanism can be substantially reinforced (mitigated) if high skill wages are more (less) dispersed compared to low skill wages.

Finally, we investigate how the two network characteristics α and φ influence the decision to acquire education, our results are summarized in lemma 2:

Lemma 2: *In both types of families more children obtain education if the network size parameter φ increases or skill homophily α decreases when $V[w_i^H]/V[w_i^L] > h/(1 - h)$ holds, that is $\partial \Phi^L(c)/\partial \varphi < 0$, $\partial \Phi^H(c)/\partial \varphi < 0$, $\partial \Phi^L(c)/\partial \alpha > 0$, $\partial \Phi^H(c)/\partial \alpha > 0$, because:*

$$\frac{\partial V[\Delta^L]}{\partial \varphi} < \frac{\partial V[\Delta^H]}{\partial \varphi} < 0 \quad \text{and} \quad \frac{\partial V[\Delta_j^L]}{\partial \alpha} > \frac{\partial V[\Delta_j^H]}{\partial \alpha} > 0$$

Proof: Appendix

Recall that parameter φ is a maximum network size in the absence of skill homophily. Lemma

2 shows that having larger networks improves the quality of information, so the estimates of the return to schooling Δ_j^k, $k = L, H$ become more accurate and less dispersed. Hence, fewer children in both types of families make mistakes. The opposite is true when the homophily parameter α is increasing. Stronger network homophily implies fewer social links between the two groups of families and reduces the precision of the skill premium estimates. In absolute terms both effects are stronger for low skill families since their networks are very asymmetric with only a few social links to high skill families making their decisions more sensitive to the network parameters (see lemma 1).

2.3   Measures of intergenerational skill dependence

In this section we derive two indicators measuring the intergenerational skill dependence: the correlation coefficient ρ and the mobility ratio m – and investigate how these measures depend on the underlying network and labour market parameters. Let ξ be a binary indicator corresponding to the education of parents, and taking value 1 for high skill parents and 0 otherwise. We know that $E[\xi] = h$ and $V[\xi] = (1-h)h$. In a similar way, let η be a binary indicator corresponding to the education of children. Based on our results from above we get the following joint distribution of ξ and η:

	$\eta = 1$	$\eta = 0$	Σ
$\xi = 1$	$hp(1 - \Phi^H(c))$	$h(1 - p(1 - \Phi^H(c)))$	h
$\xi = 0$	$(1 - h)p(1 - \Phi^L(c)))$	$(1 - h)(1 - p(1 - \Phi^L(c)))$	$1 - h$
Σ	$\tilde{h} = hp(1 - \Phi^H(c)) + (1 - h)p(1 - \Phi^L(c)))$	$1 - \tilde{h}$	1

Table 1: Joint density of the educational indicators for parents and children

Variable $\tilde{h}$ denotes an average education level in the second generation of workers, it is also a fraction of high skill workers in the second generation. Based on this joint distribution we can directly calculate the intergenerational correlation in educational choices $\rho = corr(\xi, \eta)$. We summarize our results in proposition 2:

Proposition 2: *Let $\Delta\Phi(c) = \Phi^L(c) - \Phi^H(c)$ be the gap in the probability of getting higher education in the second generation of workers, conditional on high ability and $\tilde{h} = hp(1 - \Phi^H(c)) + (1-h)p(1 - \Phi^L(c)))$ – average education level in the second generation of workers. The intergenerational correlation in educational choices ρ is then given by:*

$$\rho = p\Delta\Phi(c)\sqrt{\frac{(1-h)h}{(1-\tilde{h})\tilde{h}}} > 0 \quad \text{for} \quad 0 < \alpha < 1 \tag{1}$$

Proof: Appendix

From equation (1) we can see that the intergenerational schooling correlation is positive if social networks exhibit skill homophily ($0 < \alpha < 1$). In a special case when the economy is in a macroeconomic steady state such that the fraction of high skill workers is constant over time, $h = \tilde{h}$, we find that $\rho = p\Delta\Phi(c)$, so it is largely driven by different decisions of children in high and low skill families. Further, the theoretical transition matrix allows us to analyze the effect of

social network on the intergenerational mobility in education. Following the definition by Bauer and Riphahn (2007), we use the following measure of mobility:

$$m = \frac{Pr(\eta = 1|\xi = 1)}{Pr(\eta = 1|\xi = 0)} = \frac{1 - \Phi^H(c)}{1 - \Phi^L(c)}$$

This mobility ratio compares the chances of obtaining education in high and low skill families respectively. More precisely it means that children in high skill families are m times more likely to obtain education compared to children in low skill families. For a given $\tilde{h}$, one can say that both measures are increasing in the number of children in high skill families obtaining education $1 - \Phi^H(c)$ and decreasing in the number of children obtaining education in low skill families $1 - \Phi^L(c)$. The mobility ratio m only depends on these two variables, so the changes in m are easier to forecast. In contrast, the correlation coefficient also depends on other variables such as p, h and the endogenous share of high skill children $\tilde{h}$.

Further, we investigate the impact of p, which is the fraction of high ability chidlren, on the correlation coefficient ρ. Lemma 3 shows that ρ is increasing in p, even though the mobility ratio m doesn't depend on p.

Lemma 3: *The intergenerational correlation coefficient ρ is increasing in the share of high ability individuals in the second generation (children) p.*

Proof: Let $x = h(1 - \Phi^H(c)) + (1 - h)(1 - \Phi^L(c))$, so that $\tilde{h} = px$, then the correlation coefficient ρ can be rewritten as:

$$\rho = \Delta\Phi(c)\sqrt{(1-h)h}\frac{p}{\sqrt{(1-px)px}} = \Delta\Phi(c)\sqrt{\frac{(1-h)h}{x}}\sqrt{\frac{p}{1-px}} = \Delta\Phi(c)\sqrt{\frac{(1-h)h}{x}}\sqrt{\frac{1}{\frac{1}{p}-x}}$$

The fraction $1/p$ in the last expression is the only impact p, this shows that $\partial\rho/\partial p > 0$ $\square$.

The intuition behind lemma 3 can be described in the following way. The difference $p(\Phi^L(c) - \Phi^H(c))$ reflects a gap in the share of high ability children who make a mistake by not studying due to the poor quality of information. If the share of high ability children is increasing, meaning a higher p, then a larger number of children consider the option of acquiring education and can potentially make a mistake, so the gap $p(\Phi^L(c) - \Phi^H(c))$ is increasing, which is contributing to a stronger correlation coefficient ρ. There is also an opposing effect of p on the variance term $\tilde{h}(1 - \tilde{h})$, but we show in the proof of lemma 3 that this effect is always dominated.

2.4 Observation error

In the previous section we assumed that wages of social contacts are perfectly observable, which is a relatively strong assumption. In the present section we relax this assumption and consider a situation when wages of friends are observed with an observation error ϵ. Let ϵ_i^L be the observation error for low skill wages and ϵ_i^H – for high skill wages. We assume that the two error terms are normally distributed with a zero mean and variances $\sigma_{\epsilon L}^2$, $\sigma_{\epsilon H}^2$, respectively. Thus children in family j observe wage signals $w_i^L + \epsilon_i^L$ for low skill contacts of their parents and $w_i^H + \epsilon_i^H$ for high skill contacts of their parents. There is no error for observing the wage of the parent. The average wage of low skill friends in low skill families w_j^{LL} has the following form

and distribution:

$$w_j^{LL} = s w_j^L + (1-s)\frac{1}{n^{LL}}\sum_{i=1}^{n^{LL}}(w_i^L + \epsilon_i^L) \sim N(E[w_j^{LL}], V[w_j^{LL}])$$

with $\quad E[w_j^{LL}] = E[w_i^L] \quad$ and $\quad V[w_j^{LL}] = V[w_i^L]\Big(s^2 + \frac{(1-s)^2}{n^{LL}}\Big) + \frac{\sigma_{\epsilon L}^2(1-s)^2}{n^{LL}}$

In a similar way we obtain the distribution of the average wage of high skill friends in low skill families w_j^{LH}:

$$w_j^{LH} = \frac{1}{n^{LH}}\sum_{i=1}^{n^{LH}}(w_i^H + \epsilon_i^H) \sim N(E[w^{LH}], V[w^{LH}])$$

with $\quad E[w^{LH}] = E[w_i^H] \quad$ and $\quad V[w^{LH}] = \frac{V[w_i^H]}{n^{LH}} + \frac{\sigma_{\epsilon H}^2}{n^{LH}}$

So the skill premium estimated by children in low skill families has the following distribution:

$$\Delta_j^L \sim N\Big(E[w_i^H] - E[w_i^L], V[w_i^L]\Big(s^2 + \frac{(1-s)^2}{n^{LL}}\Big) + \frac{\sigma_{\epsilon L}^2(1-s)^2}{n^{LL}} + \frac{V[w_i^H]}{n^{LH}} + \frac{\sigma_{\epsilon H}^2}{n^{LH}}\Big)$$

In the group of high skill families, the skill premium Δ_j^H has the following distribution:

$$\Delta_j^H \sim N\Big(E[w_i^H] - E[w_i^L], V[w_i^H]\Big(s^2 + \frac{(1-s)^2}{n^{HH}}\Big) + \frac{\sigma_{\epsilon H}^2(1-s)^2}{n^{HH}} + \frac{V[w_i^L]}{n^{HL}} + \frac{\sigma_{\epsilon L}^2}{n^{HL}}\Big)$$

We can see that larger variances of the error terms, i.e. larger $\sigma_{\epsilon L}^2$ and $\sigma_{\epsilon H}^2$, increase uncertainty about friends' wages and raise variances of the estimated skill premia in both types of families. Conceptually, one problem faced by all families in the presence of the observation error is that wages of own parents are observed without errors, whereas wages of social contacts are subject to the observation error. Thus the estimation of the return to schooling suffers from a problem of heteroscedasticity, implying that a simple arithmetic average (with $s = 1/(1+n^{LL})$ for low skill families and $s = 1/(1+n^{HH})$ for high skill families) is no longer an efficient estimator of the mean. We describe this result in proposition 3:

Proposition 3: *In the presence of the observation error the variance of the return to schooling $V[\Delta_j^k]$, $k = L, H$ is minimized for $s = s^L$ in low skill families and $s = s^H$ in high skill families, where s^L and s^H coincide with a weighted least squares estimator and are given by:*

$$s^L = \frac{V[w_i^L] + \sigma_{\epsilon L}^2}{n^{LL}V[w_i^L] + V[w_i^L] + \sigma_{\epsilon L}^2} \qquad s^H = \frac{V[w_i^H] + \sigma_{\epsilon H}^2}{n^{HH}V[w_i^H] + V[w_i^H] + \sigma_{\epsilon H}^2}$$

Proof: Appendix.

Proposition 3 shows the exact weights s^L and s^H that the generation of children should assign to parental wages in order to deal with the problem of an observation error. In the absence of noise these weights are equal to $1/(1+n^{LL})$ for low skill families and $1/(1+n^{HH})$ for high skill families. In this special case the variance of the return to schooling $V[\Delta_j^k]$ is already minimized if a simple arithmetic average of all available wage observations is used for the estimation, so the estimated returns to schooling Δ_j^k, $k = L, H$ coincide with the classical OLS estimator and

are BLUE (Best Linear Unbiased Estimators). The situation is different in the presence of noise since the precision of the available wage observations is different (heteroscedasticity), thus a weighted least squares estimation should be applied giving rise to the updated weights s^L and s^H. This weighting scheme addresses the problem of heteroscedasticity and guarantees efficiency of the estimated return to schooling Δ_j^k, $k = L, H$.

In order to compare the weights for high and low skill families and also the weights with and without noise we introduce an auxiliary factor $z > 0$ such that $\sigma_{\epsilon L}^2 = zV[w_i^L]$ and $\sigma_{\epsilon H}^2 = zV[w_i^H]$. Intuitively, this means that the variance of the error term is proportional to the variance of the actual wage. Thus low (high) skill wages are observed with larger (smaller) precision if their variance is smaller (larger), that is $\sigma_{\epsilon L}^2 < \sigma_{\epsilon H}^2$ if $V[w_i^L] < V[w_i^H]$. For $z = 0$ we get the benchmark case from the previous section when friends' wages are observed without an error. This allows us to formulate the following corollary:

Corollary 1: *Let z be the factor of proportionality, such that $\sigma_{\epsilon k}^2 = zV[w_i^k]$, k=L,H. Children in high skill families should optimally assign a higher weight to the wage of their parent compared to children in low skill families since $n^{HH} = \varphi h < \varphi(1 - h) = n^{LL}$ if $h < 0.5$, $\forall z \geq 0$. If z is increasing, then children in all families should assign a higher weight to the parental wage, that is $\partial s^k / \partial z > 0$ since:*

$$s^k = \frac{1 + z}{n^{kk} + 1 + z} = \frac{1}{\frac{n^{kk}}{1+z} + 1} \quad \Rightarrow \quad s^H > s^L$$

This corollary shows that children in high skill families should optimally put a higher weight on the wage of their parents since the number of high skill contacts in their families is smaller on average than the number of low skill contacts in low skill families. Inserting the optimal weights s^k, $k = L, H$ into the equations for variances $V[\Delta_j^k]$ we get the following expressions:

$$V[\Delta_j^L] = \frac{V[w_i^H](1 + z)}{n^{LH}} + V[w_i^L]s^L \qquad V[\Delta_j^H] = V[w_i^H]s^H + \frac{V[w_i^L](1 + z)}{n^{HL}}$$

These equations show that even if the optimal weights s^k, $k = L, H$ are applied by the generation of children in order to cope with the observation error ϵ_i, it is still the case that stronger noise in the observed wages of social contacts z reduces the precision of the estimated return to schooling because $\partial V[\Delta_j^k] / \partial z > 0$, $k = L, H$.

Even though the two weights s^L and s^H yield efficient estimators of the return to schooling Δ_j^k, $k = L, H$, there could be different psychological and behavioral factors such that young individuals under- or overweight the importance of parental wages, that is $s \neq s^K$, $k = L, H$. Therefore, in the following lemma we analyze how the intergenerational schooling mobility depends on s:

Lemma 4: *The share of children obtaining education in low skill families $p(1 - \Phi^L(c))$ is increasing in s in the range $s \in [0..s^L]$, and decreasing for higher values of s. The share of children obtaining education in high skill families $p(1 - \Phi^H(c))$ is increasing in s in the range $s \in [0..s^H]$, and decreasing for higher values of s. The mobility ratio m is increasing for $s \in [s^L..s^H]$ given that $h < 0.5$.*

If s is large such that $s > s^H$ and increasing, then the fraction of educated children $1 - \Phi^k(c)$, $k = L, H$ is decreasing in s in both types of families, so the impact on the mobility indicators

m and ρ is ambiguous. However, when s is in the middle range, such that $s \in [s^L..s^H]$ and increasing, this leads to a higher mobility ratio m. The reason is that putting more weight on the parental wage in this range improves the precision of the return to schooling for children in high skill families and reduces the probability of making a mistake by not obtaining education (higher $1 - \Phi^H(c)$). In contrast, putting more weight on the parental wage for children in low skill families reduces the precision of their estimates and makes it more likely that they make mistakes by not studying.

In the next step we investigate the impact of h, which is the share of high skill families, on intergenerational mobility. This gives rise to the following proposition:

Proposition 4: *Consider the efficient estimator of the return to schooling Δ_j^k, $k = L, H$ evaluated at the optimal weight s^k, $k = L, H$. Then the share of children obtaining education in low skill families $p(1 - \Phi^L(c))$ is increasing in h. Let $(1-\alpha)\varphi^2/(\varphi + 2(1+z))^2 < V[w_i^L]/V[w_i^H]$, then there exists a unique threshold $0 < h^* < 0.5$ such that the share of children obtaining education in high skill families $p(1 - \Phi^H(c))$ is decreasing in h for $h \in [h^*..0.5]$. The mobility ratio m is also decreasing in h for $h \in [h^*..0.5]$. The threshold value h^* is given by:*

$$h^* = \frac{\varphi\sqrt{(1-\alpha)V[w_i^H]} - (1+z)\sqrt{V[w_i^L]}}{\varphi\sqrt{(1-\alpha)V[w_i^H]} + \varphi\sqrt{V[w_i^L]}}$$

Proof: Appendix

When h is increasing there are two effects on the network composition of low skill families. On the one hand, the share of low skill contacts in their networks $n^{LL} = \varphi(1-h)$ is decreasing. On the other hand, the share of high skill contacts $n^{HL} = \varphi h(1-\alpha)$ is increasing. Proposition 4 shows that this second effect always dominates for low skill families under the aforementioned assumption $V[w_i^H](1-h)/V[w_i^L]h > 1$. The reason is that the social networks of low skill families are very asymmetric (lemma 1) and gaining more high skill contacts has a stronger impact on the precision of their estimates than losing some low skill contacts. Thus a larger fraction of high skill families in the market improves the quality of information available to children in low skill families. So the share of children obtaining education in low skill families $p(1 - \Phi^L(c))$ is increasing in h.

The networks of high skill families are more balanced. The positive effect of having more high skill contacts $n^{HH} = \varphi h$ and leading to a higher precision of the estimated return to schooling $V[\Delta_j^H]$ is dominating for low values of h. However, when h is large (above h^*) losing low skill contacts from the network leads to a substantial loss in the precision and raises the variance $V[\Delta_j^H]$. Thus, fewer children in high skill families obtain education. Combining these findings proposition 4 shows that the mobility ratio should be eventually decreasing in h when h is sufficiently large.

More generally, our findings in proposition 4 suggest that the relationship between the share of high skill families h and the measures of intergenerational skill dependence should be hump-shaped. Even though the effect on the correlation coefficient is ambiguous it is even more likely to be hump-shaped than the mobility ratio since it has a variance term $h(1-h)$ in addition to the network effects described in proposition 4. We investigate the impact of h on the correlation coefficient in more details in the following numerical analysis.

3 Calibration

We calibrate the model using data from the German Socio-Economic Panel (SOEP). The purpose of calibrating the model is twofold. The first goal is to provide a range of quantitative estimates of the correlation coefficient that can potentially be attributed to the mechanism described in our model, that is the impact of parental networks on educational decisions of the young generation via the expectation channel. The second goal is to perform comparative statics with respect to parameters, whose theoretical effect on the correlation coefficient and the mobility ratio is ambiguous.

Our empirical data (SOEP) is a representative longitudinal household panel and approximately 15000 households and 30000 individuals participate in the survey every year (details can be found in Goebel et al. (2019)). For calibrating the social networks we use the SOEP survey from 2016 which includes a "Family and Social Networks" module. More specifically, respondents were asked to provide some information about three friends or acquaintances. The inclusion of this module makes the 2016 SOEP cross section suitable for this part of our calibration. Additionally, the SOEP provides biography information on the surveyed individuals including own and parental education.

We assume the second generation to be the survey respondents in the age range 21-65. We exclude the individuals for whom we do not have information on secondary education. This leads to a sample size for this part of the calibration of $N = 14705$. We are interested in the decision of the second generation individuals whether to become high-skilled or not. In order to explain which educational outcome corresponds to a high-skilled person we give a brief overview of the German school system and describe our classification choice.

The German secondary education system is based on parallel tracks. After elementary school children are selected into either lower secondary school (Hauptschule), intermediate secondary school (Realschule) or upper secondary school (Gymnasium) (Schindler, 2017). Alternatively, students can attend a comprehensive school (Gesamtschule) which provides more flexibility and allows students to switch between tracks (Cooke, 2003). Attending upper secondary school (Gymnasium) and obtaining upper secondary school degree certificate (Abitur) is the prerequisite for starting a tertiary degree (Lohmann and Witzke, 2011; Witte and Kalleberg, 1995). However, not all students who attain Abitur proceed to enrol in university. Also, there are options for students with intermediate secondary school degree to build upon and qualify for enrollment in a tertiary degree (Witte and Kalleberg, 1995). We define as high-skilled an individual who has obtained an upper secondary school degree certificate (Abitur). Further, our cross tabulations show that some respondents obtained a tertiary degree or were enrolled at a university or university of applied science at the time of the survey, but did not have Abitur. This group is also defined as high-skilled.

The "Family and Social Networks" module in the 2016 SOEP survey allows us to construct a social network of the respondents and estimate the relative network homophily α. Specifically, the survey participants were asked to consider three friends or acquaintances. The exact formulation of the question is: "Please think of three people outside of your household who are important for you, personally. They can be relatives or non-relatives". The respondents are then asked to provide information about their social connections including the highest school degree

attained. We assume that a high-skilled contact is a person who has "Abitur / Hochschulreife / Fachhochschulreife" which corresponds to Upper Secondary School Degree and eligibility to enrol in a tertiary degree. We do not have information on the tertiary education of the friends and acquaintances. Table 2 shows the skill distribution of the three friends and acquaintances conditional on the skill level of the survey respondents. For generating the network, we consider only those who provided information about the school certificate of all three of their friends. This reduces the sample size to $N = 10304$.

| | Number of high skill friends | | | | |
Skill level	0	1	2	3	Total
Low	3134	1649	810	291	5884
High	435	968	1442	1575	4420
Total	3569	2617	2252	1866	10304

Table 2: Number of high skill friends conditional on own skill level. Source: 2016 German Socio-Economic panel, own calculations.

Table 2 shows that the fraction of high-skilled respondents in the 2016 sample with complete information about their social contacts is 42.9% meaning that $h_{2016} = 0.429$. Focusing first on the networks of low-skilled interviewees, we observe a total of 17652 connection since each respondent is asked to name three friends or acquaintances. Out of those connections, 13510 ($= 3 \times 3134 + 2 \times 1649 + 810$) are with other low-skilled individuals. This means that γ_L, the fraction of contacts of type L in the network is $\gamma_L = \frac{13510}{17652} \approx 0.765$. Similarly, there are 4420 high-skilled interviewees which corresponds to 13260 connections. 8577 ($= 3 \times 1575 + 2 \times 1442 + 968$) of those connections point to other high-skilled individuals, which implies $\gamma_H = \frac{8577}{13260} \approx 0.647$. If the formation of social links did not exhibit skill homophily ($\alpha = 0$), we would expect that the fraction of same type contacts, γ_L and γ_H equal the fractions of low- and high-skilled individuals in the network, respectively. We see that this is not the case. While, 42.9% of the individuals in the sample are high-skilled, 64.7% of their contacts are also high-skilled. On the other hand, 57.1% of the respondents are low-skilled but 76.5% of their contacts are also low-skilled.

In order to calculate the relative network homophily α, we use the fact that the number of same skill contacts is $\varphi((1-h)^2 + h^2)$ while the total number of contacts is $\varphi(1 - 2\alpha h(1-h))$. This implies:

$$\frac{\varphi((1-h)^2 + h^2)}{\varphi(1 - 2\alpha h(1-h))} = \frac{8577 + 13510}{13260 + 17652}$$

Based on table 2 we have $h_{2016} = 0.429$. Inserting this value, we find the relative network homophily $\alpha = 0.584$.

In order to check whether the estimated homophily parameter α is stable over time we use two earlier waves of the SOEP (2011 and 2006) containing information about social networks[7]. We find that $\alpha = 0.598$ in the 2011 SOEP survey and $\alpha = 0.595$ in the 2006 survey. Table 3 shows comparison of the networks based on the three waves of the SOEP survey. We see that the estimate for the relative network homophily is robust and does not change substantially over time. This finding suggests that network homophily is persistent across generations.

To calibrate the relevant labour market parameters, we use the more recent 2018 wave of the

---

[7]The module "Family and Social Networks" is included into the SOEP questionnaire once in five years starting from 2006

SOEP survey year	Sample size	γ_L	γ_H	h	α
2016	10304	0.765	0.647	0.429	0.584
2011	8211	0.787	0.638	0.410	0.598
2006	10927	0.814	0.602	0.389	0.595

Table 3: Homophily estimates. 2016, 2011 and 2006 SOEP surveys, own calculations.

	Mean	SD	Var	Min	Max	N
Second generation						
Age	44.93	12.17	148.2	21	65	15932
High-skilled (0/1)	0.378	0.485	0.235	0	1	15932
Net income (log) (HS)	7.708	0.606	0.367	2.3	9.6	4376
Net income (log) (LS)	7.345	0.543	0.294	4.1	9.6	6509
First generation						
High-skilled mother (0/1)	0.135	0.342	0.117	0	1	14679
High-skilled father (0/1)	0.201	0.401	0.161	0	1	14679
High-skilled parent (0/1)	0.248	0.432	0.187	0	1	14679

Table 4: Summary statistics. Source: 2018 SOEP, own calculations.

SOEP survey which provides us with a larger sample but does not include information about social connections. Again, we consider survey respondents in the age range 21-65 and exclude those for whom we do not have information about their secondary education. This leads to a sample size of 15932 individuals, 37.8% of whom are high-skilled, so we know that $\tilde{h} = 0.378$. SOEP 2018 has detailed information about parental education. The definition of high-skilled parent is consistent with that of a high-skilled child. We assume that the child has a high skill parent if either one of the parents is high-skilled. If information on one of the parents is missing, then this parent is assumed to be low-skilled and we consider only the skill level of the second parent. The lower panel of table 4 further shows that 24.8% ($h = 0.248$) of the children have a high skill parent where 20.1% of the fathers are high-skilled compared to only 13.5% of the mothers. Our choice of parameters is summarized in table 6 in the end of this section.

Recall that networks in the model are driven by a combination of three parameters α, h, φ. So the remaining network parameter to be determined is the maximum network size φ. Since our data does not include information about the network size we rely on the common estimates in the literature. For example, Cingano and Rosolia (2012) report that the number of social connections between individuals in Italy is about 32. Glitz (2017) reports a similar number for Germany with approximately 43 social contacts. So we set $\varphi = 38.4$ targeting the average network size $n = \varphi[1 - 2\alpha h(1-h)] = 30$. Given a full vector of network parameters $\alpha = 0.584, h = 0.248, \varphi = 38.4$ we can characterize the structure of social networks in the model: $n^{LH} = 3.95$, $n^{HH} = 9.51$, $n^{HL} = 11.99$ and $n^{LL} = 28.84$. The corresponding Poisson densities are illustrated on figure 1. Inline with lemma 1 the distribution of social contacts is more balanced in high skill families since the numbers of high and low skill contacts are similar, while it is more extreme in low skill families, where the number of low skill contacts is much higher than the number of high skill contacts.

Next, we estimate the first and second moments of the wage distributions of low- and high-skilled workers. The survey respondents are asked how much they earned from work *last* month.
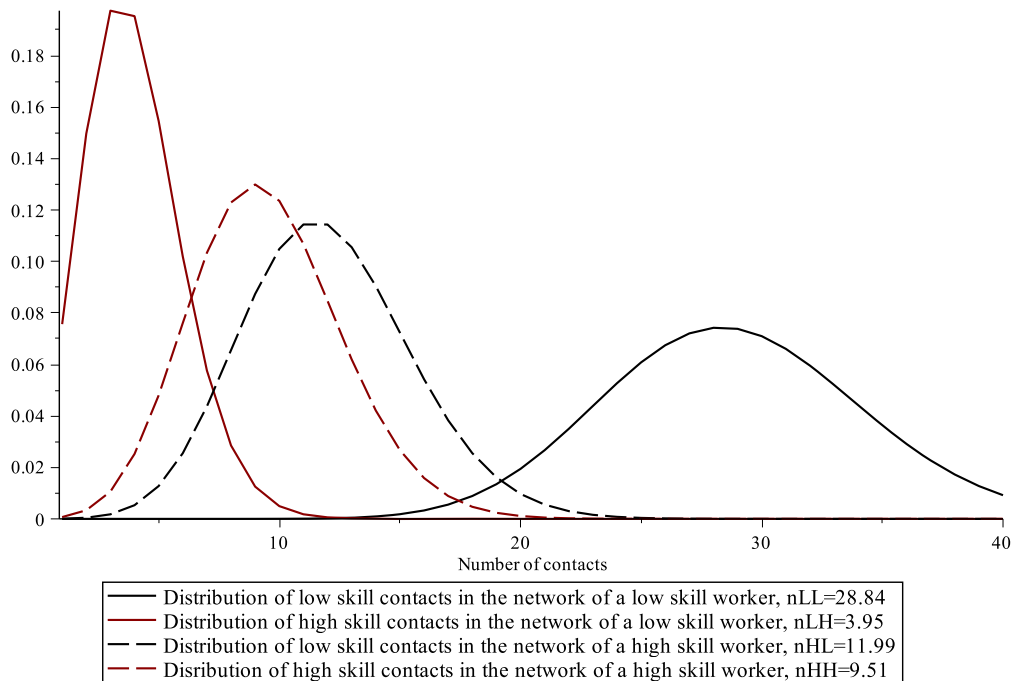
Figure 1: Network densities with effects of skill homophily and lower fraction of high skill contacts

This includes both wages earned as an employee as well as income from self-employment. We consider only those who declare that they are full- or part-time employed. Figure 2 shows the kernel density estimates of the wage distributions conditional on skill level. We drop wages above the 99.9^{th} percentile of the overall income distribution in order to exclude outliers from the calibration. Table 4 provides summary statistics by skill group. We observe that the wage distribution of high-skilled workers has a higher mean and a higher variance. The average net monthly log income of high-skilled workers is $E[w_i^H] = 7.708$ compared to $E[w_i^L] = 7.345$ for low-skilled workers[8] which implies a skill premium of 0.363. Further, the wages of high-skilled workers are more dispersed compared to the wages of low-skilled workers such that $V[w_i^H] = 0.367 > V[w_i^L] = 0.294$ (see figure 2). These parameter values satisfy conditions derived in proposition 1.

Several remaining parameters are associated with the observation noise ϵ_i^j, these parameters are z, s^H, s^L. Given that the variance of the error term $\sigma_{\epsilon j}^2$ is not observable, we rely on the fact the standard deviation of the error term $\sigma_{\epsilon j}$ can be interpreted as a %-deviation from the mean log-wage and calibrate the model to match a $20 - 25\%$ deviation. This target can be achieved by setting $z = 0.15$, and implies that $\sigma_{\epsilon L} = 0.21$ and $\sigma_{\epsilon H} = 0.24$. Intuitively, this means that if the real unobservable wage of a high skill social contact is 2225 EUR (i.e. the average high skill wage), then the friends of this person expect it to be in the range [1691..2759] with a 68% probability. Given $z = 0.15$ we use equations derived in proposition 3 and calculate the optimal weights $s^L = 0.038$ and $s^H = 0.107$. Inline with the results in corollary 1 these values are above $1/(n^{LL} + 1) = 0.033$ and $1/(n^{HH} + 1) = 0.095$ respectively. So children should optimally put a higher weight on the wage of their parent if it is the only observation without noise.

---

[8]This corresponds to mean net incomes of 2225 Euros and 1549 Euros for high- and low-skilled workers, respectively.
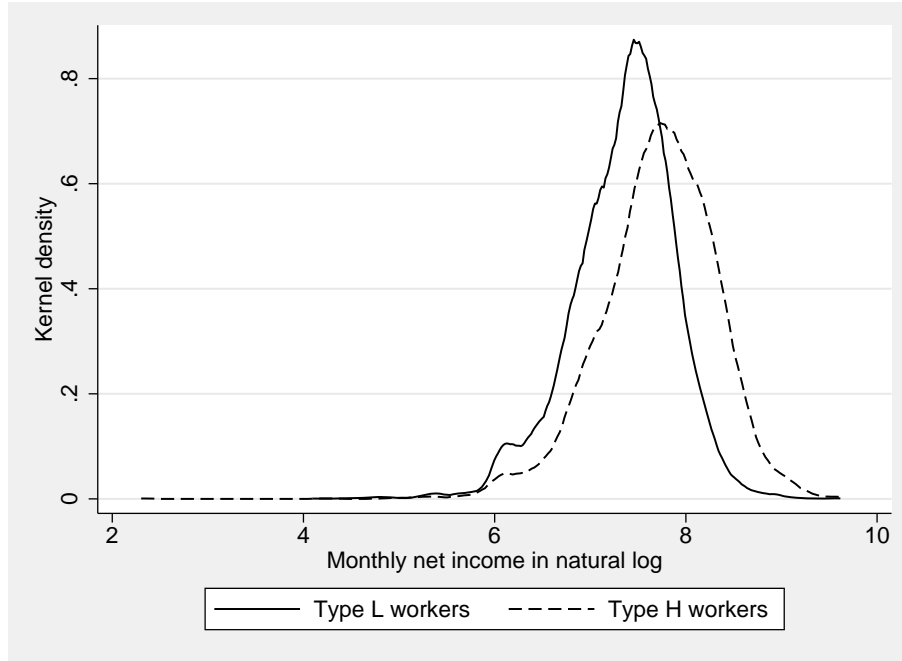
Figure 2: Kernel density estimates of (log) net income of low- and high-skilled workers in part and full time employment. The top 0.1% of net incomes are dropped from the estimation (this corresponds to net income of above 15000 Euros)

Based on the calibrated parameters we can find variances of the two skill premiums $V[\Delta_j^H] = 0.068$ and $V[\Delta_j^L] = 0.118$. The corresponding distributions of the estimated skill premiums $\Phi^L(.)$ and $\Phi^H(.)$ are illustrated on figure 3. Note that both distributions have the same mean $E[w_i^H] - E[w_i^L] = 0.363$, which is the average income gap between high and low skill workers. Yet, the expectations of children in low skill families are more dispersed compared to the expectations of children in high skill families. The right panel of figure 3 shows that for every cost $c < E[w_i^H] - E[w_i^L] = 0.363$ children in high skill families are more likely to study than children in low skill families inline with proposition 2.
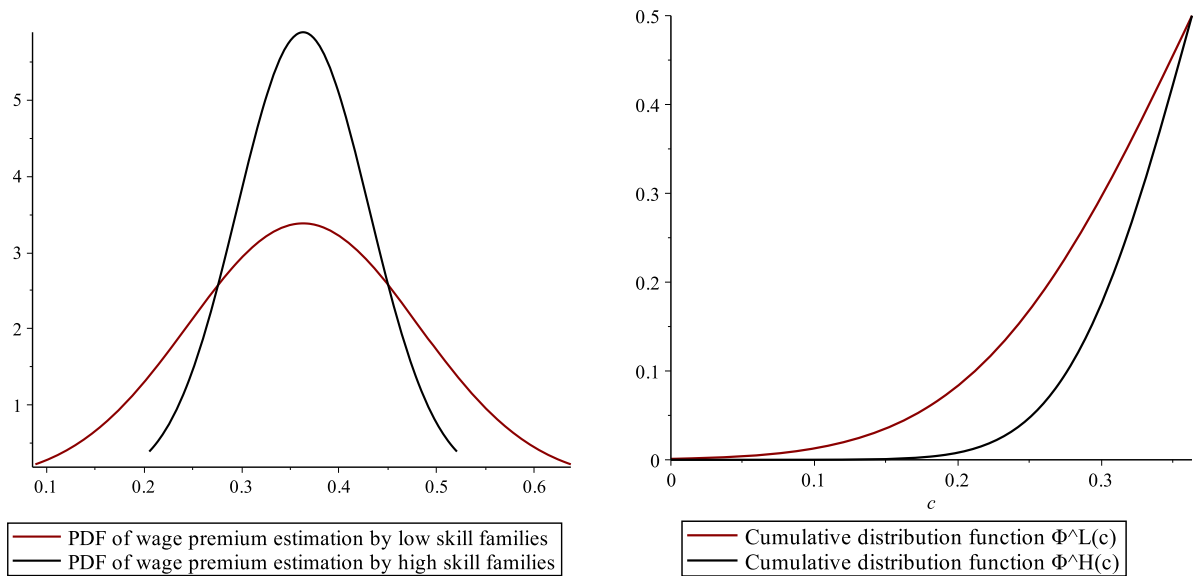

Figure 3: PDF and CDF of wage premium estimations by low and high skill families

23

The last parameter p is the probability that the child has high abilities. We calibrate this parameter by targeting the fraction of high skill individuals in the 2018 SOEP sample, that is $\tilde{h} = 0.378$, producing $p = 0.474$.

Finally, we estimate the intergenerational skill dependence using the educational transition matrix shown in table 5. The values on the main diagonal are well above 50% which points towards a low level of educational mobility. 71% of children born into high-skilled families became high-skilled themselves, compared to only 29% of children born in low-skilled families. Using table 5 and the result from Proposition 2, we estimate the intergenerational correlation in education $\rho = 0.375$. We use this parameter in order to have a reference point and answer the question how much of the empirically observed correlation coefficient can be attributed to the network-driven skill transmission mechanism described by the model.

	$\eta = 1$	$\eta = 0$	Σ
$\xi = 1$	2591	1055	3646
	(0.71)	(0.29)	
$\xi = 0$	3161	7872	11033
	(0.29)	(0.71)	
Σ	5752	8927	14679
	(0.39)	(0.61)	

Table 5: Transition matrix. Fractions in parentheses. ξ: skill level of the parent; η: skill level of the children. Source: 2018 German Socio-Economic panel, own calculations.

The corresponding mobility ratio m is given by:

$$m = \frac{Pr(\eta = 1|\xi = 1)}{Pr(\eta = 1|\xi = 0)} = 0.71/0.29 = 2.48$$

which shows the probability that a child born into a high-skilled family becomes high-skilled relative to a child born into a low-skilled family. There is a two and half time difference in the two probabilities.

4 Comparative statics

In this section, we provide comparative statics results and explore how the intergenerational schooling correlation and educational mobility depend on the structure of social networks of parents and labour market parameters. Specifically, we investigate the impact of the cost parameter c, the share of high skill families h, network parameters φ and α and the main uncertainty parameter z. This analysis complements our previous results in situations where the theoretical effect of a given parameter is ambiguous. Moreover, it allows us to make conclusions about the elasticity of the two dependence measures ρ and m with respect to the key parameters of the model. All our numerical results refer to the modified model with uncertain wage observations.

First, we start by analyzing the impact of the education cost c by varying it in the range [0..0.36]. The left panel of figure 4 illustrates the effect of changes on the intergenerational schooling correlation for the different values of $p = [0.37, 0.47, 0.57]$. We find that the impact of the cost parameter on the correlation coefficient is non-monotone and its magnitude is limited from above. The same result is observed for the mobility ratio m on the left of figure 5. This

Table 6: Baseline parameters

Parameter	Value	Description and explanation
α	0.584	Relative network homophily, SOEP 2016
h	0.248	Fraction of high skill parents, SOEP 2018
φ	38.35	Targeting the average network size $n = 30$. Consistent with in Cingano and Rosolia (2012), Glitz (2017) and Fontaine (2008).
$E[w_i^H]$	7.708	Average wage of high skill workers (log), SOEP 2018
$V[w_i^H]$	0.367	Variance of high skill wages, SOEP 2018
$E[w_i^L]$	7.345	Average wage of low skill workers (log), SOEP 2018
$V[w_i^L]$	0.294	Variance of low skill wages, SOEP 2018
z	0.150	Targeting the standard deviation of the error term $20-25\%$
s^H	0.107	Weight to the own parents wages for high skill workers. WLS
s^L	0.038	Weight to the own parents wages for low skill workers. WLS
p	0.474	Probability of being high ability for children. Targeting the fraction of high skill children $\tilde{h} = 0.378$, SOEP 2018
c	0.278	Cost of education. Corresponds to the upper bound of intergenerational correlation in schooling

is intuitive: if the cost is very low, then almost all high ability children in high and low skill families acquire education and differences in the expected skill premium do not play any role for their decisions. If the cost is high and close to the skill premium, then roughly a half of high ability children acquire education irrespective of their parental background and networks. This shows that parental networks have the strongest impact on the decisions of children at intermediate values of the educational cost. For example, ρ reaches its maximum value of 0.055 at the education cost $c = 0.278$ when $p = 0.474$ which is the benchmark case. Given that our data does not include information about the cost c, moreover, this cost is unobservable since it includes monetary and non-monetary components (e.g. effort) we interpret our result as an upper bound of the correlation coefficient. We believe that this approach is more informative than an alternative point estimate based on ad-hoc assumptions about the cost. Given that the empirically observed correlation coefficient is equal to 0.375, we conclude that parental networks and their impact on expectations can possibly explain up to 14.7% of the intergenerational correlation coefficient (that is 0.055/0.375). The upper bound of the mobility ratio is 1.177, which means that the probability of acquiring higher education is almost 1.2 times higher for children in high skill families compared to children in low skill families. Note that these estimates refer alone to the impact of parental networks on expectations and do not include other factors.

The right panel of figure 4 shows that there is a non-monotone relationship between the intergenerational schooling correlation and the share of high skill families h. The same is true for the impact on the mobility ratio on figure 5. The reason is that a higher fraction of high skill parents reduces the informational disadvantage faced by children in low skill families, so that more children in these families acquire education inline with proposition 4. But the effect on children in high skill families is ambiguous, on the one hand, there is lower uncertainty about

high skill wages, but on the other hand, there is higher uncertainty about low skill wages. The first effect dominates when h is small, implying a steep increase in $1 - \Phi^H(c)$, whereas the second effect dominates when h is sufficiently large, implying a moderate reduction in $1 - \Phi^H(c)$. This leads to a hump-shaped relationship between the share of high skill parents and the two mobility measures ρ and m, whereby for ρ the effect is amplified by the additional term $\sqrt{h(1-h)}$. We find that the pivotal point h^* described in proposition 4 is equal to 0.35, however, both mobility measures achieve a maximum for lower values of h.

Summarizing, on a theoretical level we find that the probability of obtaining higher education $p(1 - \Phi^L(c))$ is increasing and concave implying a falling marginal gain for children in low skill families. This is also the case in high skill families since $p(1 - \Phi^H(c))$ is sharply increasing and concave for low values of h but it starts declining for larger values of h. Yet, numerically we find that the decline is rather small. We take this as a main prediction of the model and test it empirically in section 5.

Figure 4: Comparative statics with respect to the educational cost c and the fraction of high skill parents h for different values of p

The right panels of figures 6 and 7 illustrate the impact of the skill homophily parameter α. Even though the quality of information is worsening with α in both types of families and fewer children obtain education (lemma 2), the negative consequences are stronger for low skill families when α is small. The reason is that their networks are less balanced (lemma 1), so a stronger group segregation (i.e. higher α) rapidly reduces the small number of valuable high skill contacts in low skill families. However, as $\alpha \to 1$ both types of families become "infinitely uncertain" about the market skill premium leading to a reversed effect of α. In the limiting case $\alpha \to 1$ the impact of the family background becomes irrelevant since:

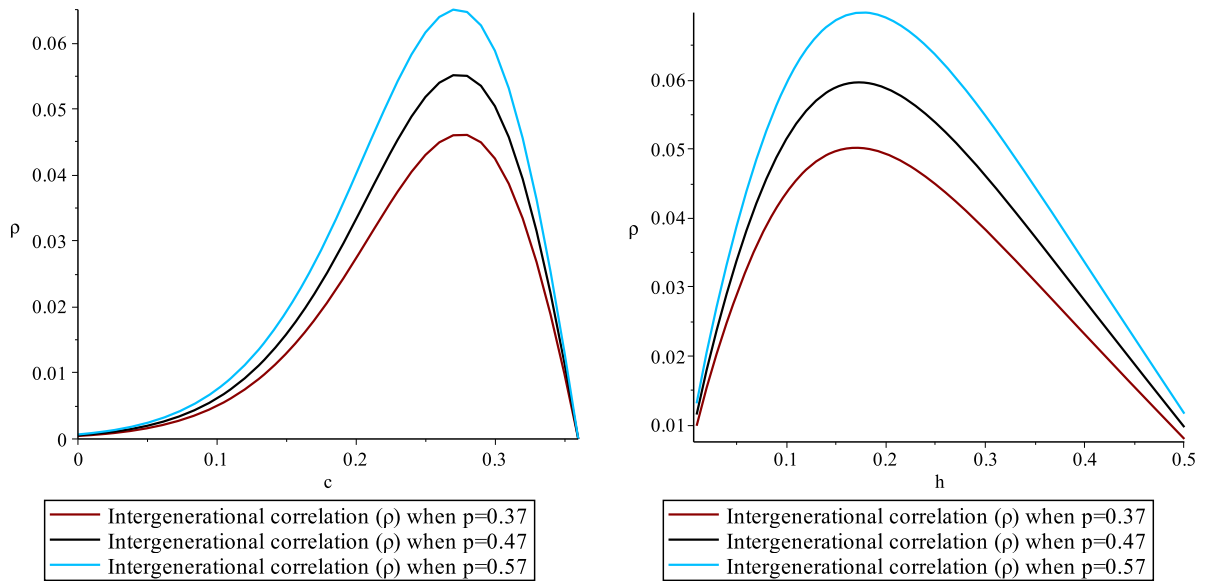$$lim_{\alpha \to 1} p(1 - \Phi^k(c)) = 0.5p \quad \lim_{\alpha \to 1} \rho = 0 \quad \lim_{\alpha \to 1} m = 1$$
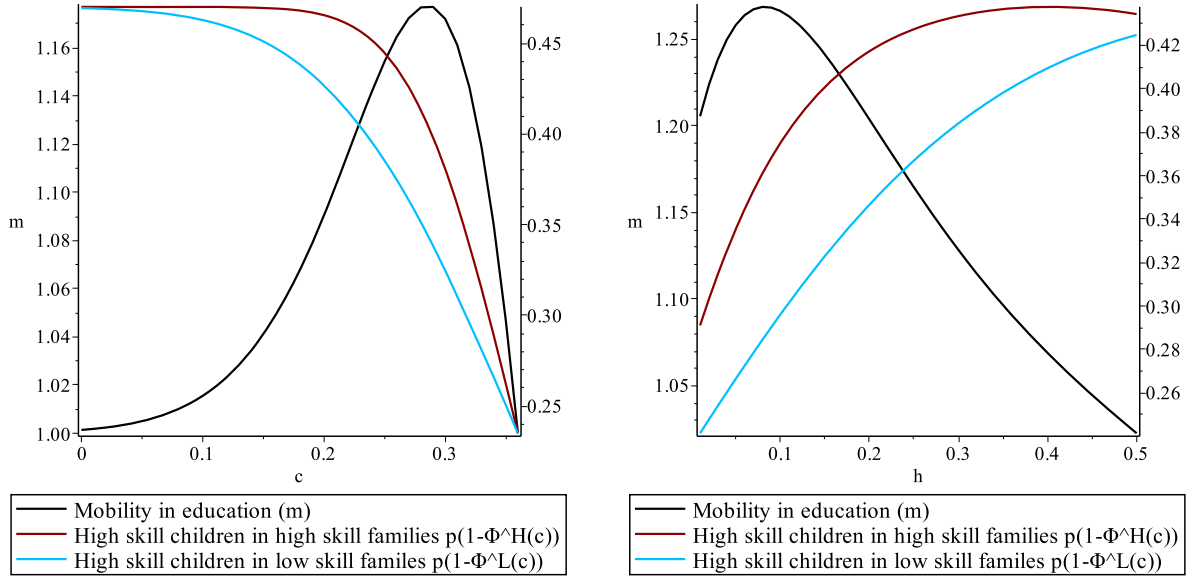
26

Figure 5: Comparative statics with respect to the educational cost c and the fraction of high skill parents h

One interesting observation from the right panels of both figures is that there exists a lower bound of the correlation coefficient ρ and mobility ratio m, when there is no network homophily: $\alpha = 0$. Comparing variances of the two skill premiums for $\alpha = 0$, we get:

$$V[\Delta_j^L] - V[\Delta_j^H] = \frac{(1+z)^2}{\varphi} \left(\frac{V[w_i^H]}{h(\varphi h + 1 + z)} - \frac{V[w_i^L]}{(1-h)(\varphi(1-h) + 1 + z)} \right)$$

This expression is strictly positive if the weighted variance ratio $V[w_i^H](1-h)/(V[w_i^L]h)$ is above 1, which holds in our data. This leads to a situation when the number of high skill children acquiring education in high skill families $p(1 - \Phi^H(c)) = 0.449$ is larger than the number of children acquiring education in low skill families $p(1 - \Phi^L(c)) = 0.443$, so that $\rho > 0$ and $m > 1$. Intuitively, this situation arises because some children have high skill parents and, therefore, one high skill contact more on average (and one low skill contact less) than children in low skill families. Given that information about high skill wages is more valuable since high skill wages are more dispersed and the number of high skill people in the market is small, this asymmetry benefits children in high skill families and leads to $\rho > 0$ even if parental networks are identical and there is no skill homophily, that is $\alpha = 0$.

Further, we vary the size parameter φ between 20 and 60, so the average network size n changes between 15 and 50. We can see that both variables have an inverse U-shape influence on ρ and m. As demonstrated in lemma 2, a higher size φ reduces uncertainty about the market skill premium in both types of families. For high skill families this effect is stronger at low values of φ, because their social networks are smaller on average (see lemma 1) compared to low skill families. Thus the marginal gain of having larger social networks is stronger for high skill families leading to an increasing correlation coefficient ρ at low values of φ. However, as networks get large the marginal gain of high skill families is diminishing and falls below that for low skill families, leading to a reverse relationship between φ and ρ.

27
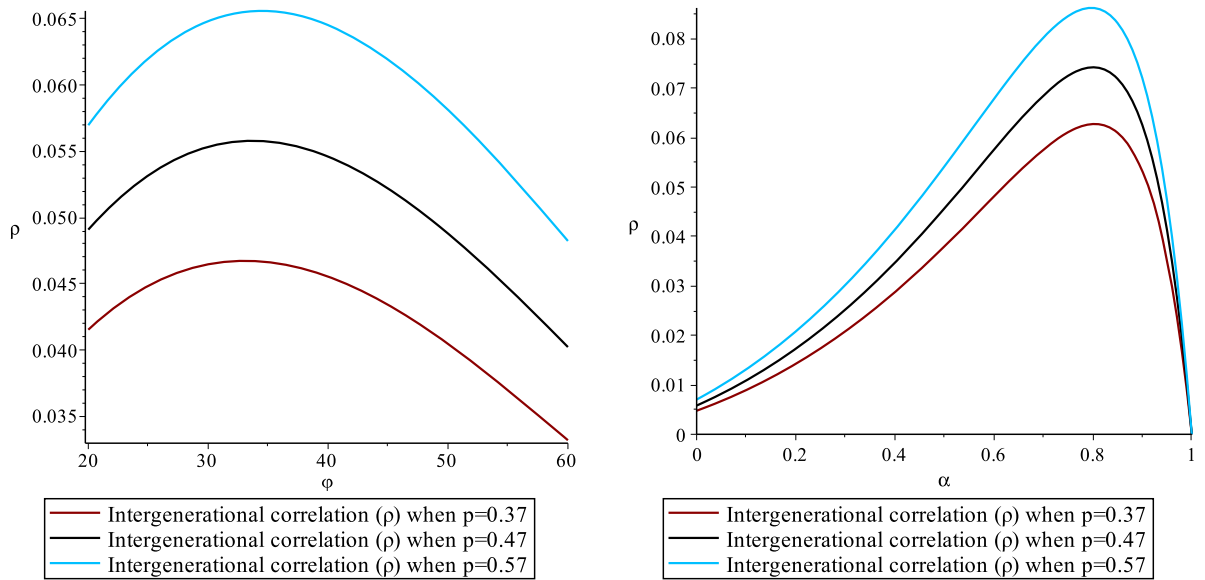
Figure 6: Comparative statics with respect to parameter φ and the skill homophily α for different values of p

Note, that our numerical results allow us to make a number of quantitative statements concerning the sensitivity/elasticity of the two measures of intergenerational mobility ρ and m to the considered parameters. In particular, we can see the effect of the network size φ is moderate, however, both measures ρ and m are very sensitive to the educational cost c, the share of high skill families h and the degree of skill homophily α.
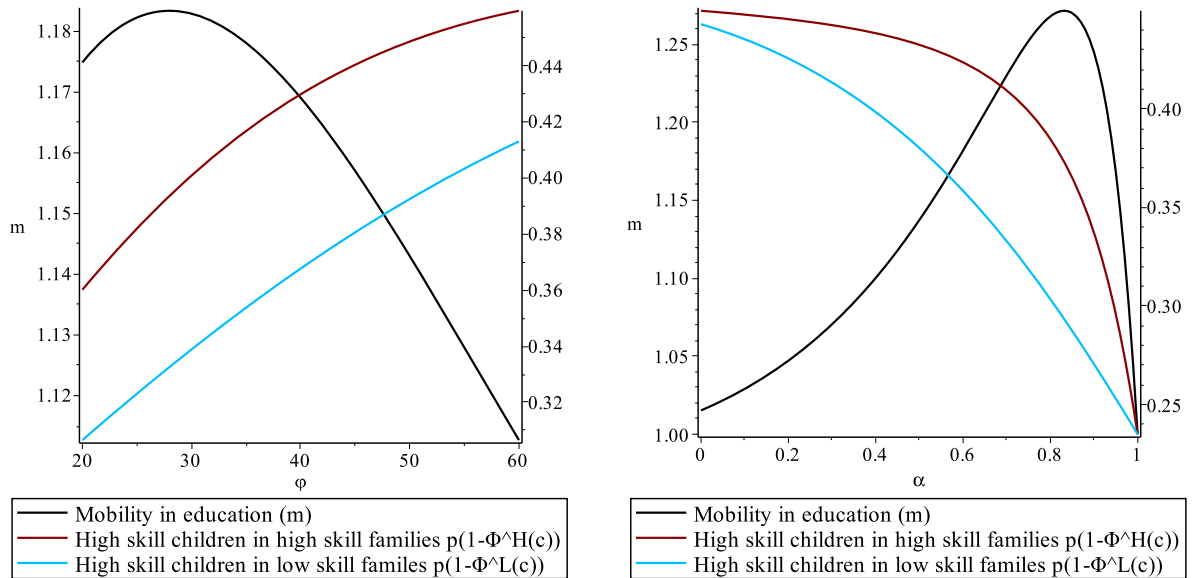
Figure 7: Comparative statics with respect to the parameter φ and the skill homophily α

Finally, we analyze the impact of the observation error ϵ_i^K on both measures of intergenerational mobility ρ and m. We achieve this goal by varying parameter z in the range $[0..1]$. Figure 8 shows that both types of families suffer from higher uncertainty when z is small and increasing, implying that fewer children obtain education. Yet, low skill families are stronger affected since their social networks are more asymmetric (lemma 1), so both measures ρ and m are increas-

ing initially. However, differences in the available information disappear as $z \to \infty$, meaning that both types of families become "infinitely uncertain" about the skill premium leading to a reversed effect of z on ρ and m for large values of z. For the limiting case we find that:

$$\lim_{z \to \infty} p(1 - \Phi^k(c)) = 0.5p \quad \lim_{z \to \infty} \rho = 0 \quad \lim_{z \to \infty} m = 1$$

Quantitatively, the effect of the observation error is rather small compared to all other parameters, that is both measures ρ and m are not sensitive to the variance of the error term. The reason is that social networks of parents in our benchmark specification are rather large, so the estimates of the skill premium are relatively precise despite the presence of noise.
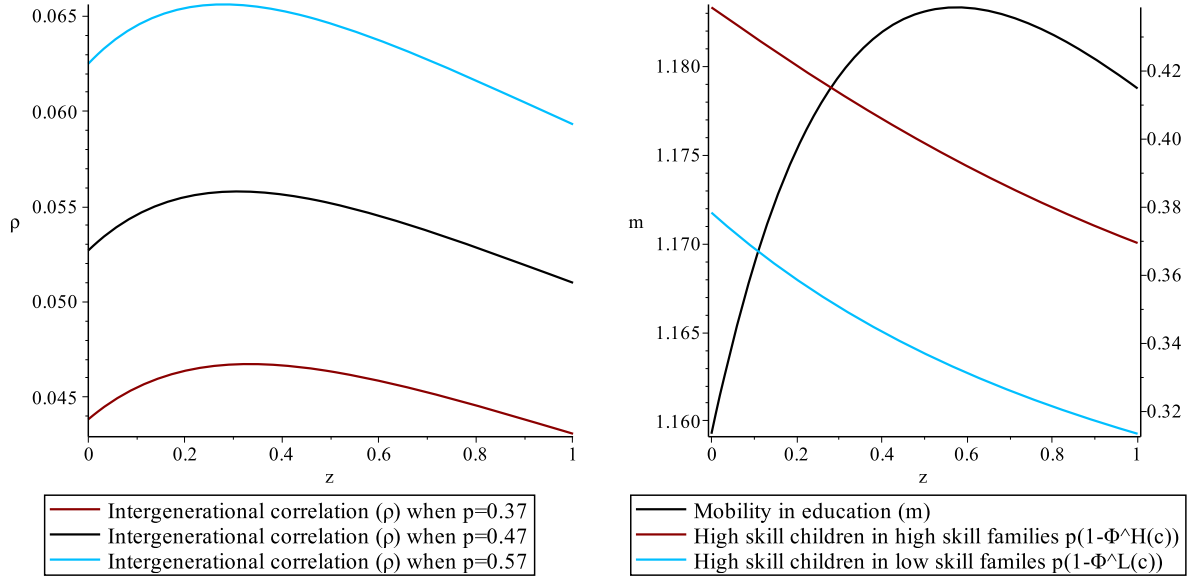
Figure 8: Comparative statics with respect to z: Left panel: the intergenerational correlation for different values of p. Right panel: mobility

Concluding this section, we can see that parameters h and φ have a similar qualitative impact on both dependence measures in the sense that both types of families are gaining from more precise information, however, high skill families gain more at low values of h and φ because their social networks are smaller on average. Considering the impact of variables α and z, both types of families suffer from higher uncertainty when α and z are small and increasing, but the negative impact is stronger for low skill families since their networks are more asymmetric. In all considered cases this leads to an inverse U-shape relationship between the respective parameter and the two measures ρ and m. Quantitatively, the effect of variables z and φ is small to moderate, because social networks of both types of families are sufficiently large on average, but the impact of variables h and α is strong.

5 Empirical Analysis

In this section we use SOEP data to test the main prediction of the theoretical model that the quality of parental social networks can have an impact on the schooling decisions of children. In our empirical analysis we follow the approach described in Plug et al. (2018) with the exception

that the outcome variable of interest is a schooling level, whereas it is the earnings score of a child in Plug et al. (2018). We follow the same definition of being "high skilled" for the child as in the calibration stage and use a logit model to predict the probability of obtaining higher education $P\{\eta = 1\}$. Our theory predicts that a higher number of high skill individuals in the economy h is positively associated with a share of high skill friends in the networks of parents in family i, let it be denoted by H_{it}[9]. Moreover, the probability of obtaining higher education for the child should be increasing in the share of high skill family friends H_{it} with a falling marginal gain. For high skill families the marginal gain may even become negative but numerically the decline in the probability is small (proposition 4 and figure 5, right panel). We capture this non-monotonicity by including variables H_i and H_{it}^2 in the regression equation. We complement these variables with the years of schooling of both parents, demographic control variables including gender, age, age squared, immigration background and the presence of siblings in the family as well as indicators of cognitive and non-cognitive abilities of parents and children and time fixed effects.

As already mentioned above, information about social networks is collected in the years 2006, 2011 and 2016. Based on these waves we calculated the shares of high skill friends separately for mothers and fathers. Table 7 contains a list of main variables with their descriptive statistics. In addition, we constructed a combined variable H_{it}, which is the share of high skill friends of a family (both parents). Recall that both parents are asked about the education of their three best friends, this gives rise to the information about six best friends of both parents[10] (living outside the household). Hence, our main variable of interest H_{it} is calculated as a total number of high skill friends of both parents divided by six. The distribution of this variable is presented on figure 9. It shows that 34% of families do not have any high skill friends, whereas 7.6% of families report all six friends being high skilled.
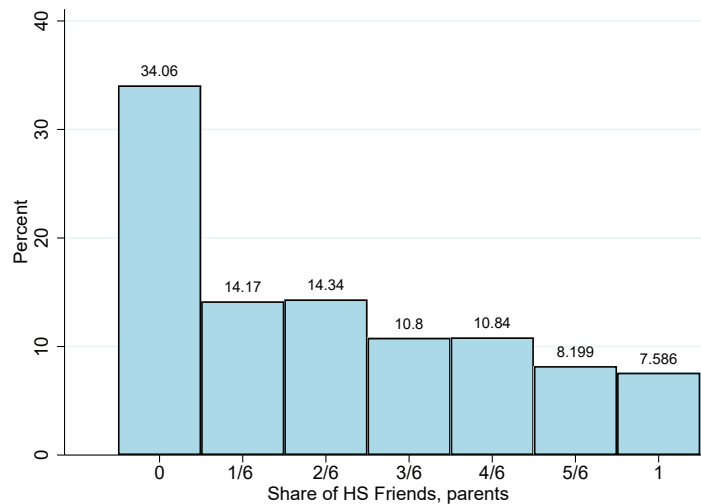

Figure 9: Share of high skill friends in the family H_{it}

---

[9]Recall that $H_L = 1 - \gamma_L$ and $H_H = \gamma_H$, both increasing in h.

[10]We can not exclude the overlap with the same person being the best friend of both parents. Yet, a strong gender bias in the networks of best friends suggests that the overlap is likely to be small.

	Mean	SD	Min	Max	N
Share of HS Friends, Mother	0.320	0.357	0	1	7406
Share of HS Friends, Father	0.336	0.370	0	1	5843
Share of HS Friends, parents	0.342	0.333	0	1	4574
Years of Schooling Mother (0-13)	9.241	2.434	0	13	51262
Years of Schooling Father (0-13)	9.548	2.393	0	13	49546
Cognitive Ability	9.035	4.384	0	28	23423
Cognitive Ability Mother	7.719	4.054	0	25	3582
Cognitive Ability Father	7.680	4.193	0	23	2681
Ext. Locus of Control I	0.463	0.253	0	1	44867
Ext. Locus of Control II	0.689	0.192	0	1	45166
Age	42.139	13.026	16	66	64814
Male gender	0.464	0.499	0	1	64814
Siblings	0.889	0.315	0	1	63109
First Generation Immigrant	0.193	0.395	0	1	64814
Second Generation Immigrant	0.057	0.233	0	1	64814

Table 7: Descriptive statistics, SOEP 2006, 2011, 2016

As highlighted in Plug et al. (2018), the effect of parental networks may be overestimated if there are (omitted) factors that enable parents to form friendships with high skill contacts and help their children in obtaining higher education. The key candidate here is cognitive ability of the parent, since high ability parents are likely to have high ability children and form links to other high ability individuals. In order to address this argument and avoid/reduce an upward bias we include proxy variables for cognitive abilities of parents and children into the regression. SOEP waves 2006, 2012 and 2016 contain information about an ultra-short symbol digit test (for a detailed description see Lang et al. (2007)). The symbol digit test (SDT) requires individuals to match numbers with graphical symbols according to a given conversion table. The task is accomplished by pressing numbers on the PC keyboard as quickly as possible. The number of correct answers is recorded after 30, 60 and 90 seconds and provides an estimate of the respondent's perceptual information-processing speed. This test reveals innate abilities of the individual and is related to the speed of solving new tasks. It was already used in other influential studies as a proxy for cognitive abilities (e.g. Anger and Heineck (2010)) and was shown to produce outcomes which are sufficiently correlated with test scores of more comprehensive intelligence tests (Lang et al. (2007)). We matched the test performance of the respondents in the year 2012 to their personal ID numbers in 2011 since cognitive abilities of the individual are mostly inborn and it is unlikely that they changed substantially within a year. Further, we use information about the SDT performance of the individual in the first 30 seconds, since this initial period is associated with the highest variance in the performance. We deleted several extreme outliers. Figure 10 shows the distribution of test results in the generation of children. The average value is 9 correct answers in 30 seconds, but there is a substantial variation with a standard deviation equal to 4.3. The distributions of cognitive abilities for parents have similar shapes but the mean values are lower. This is because parents are older at the time of
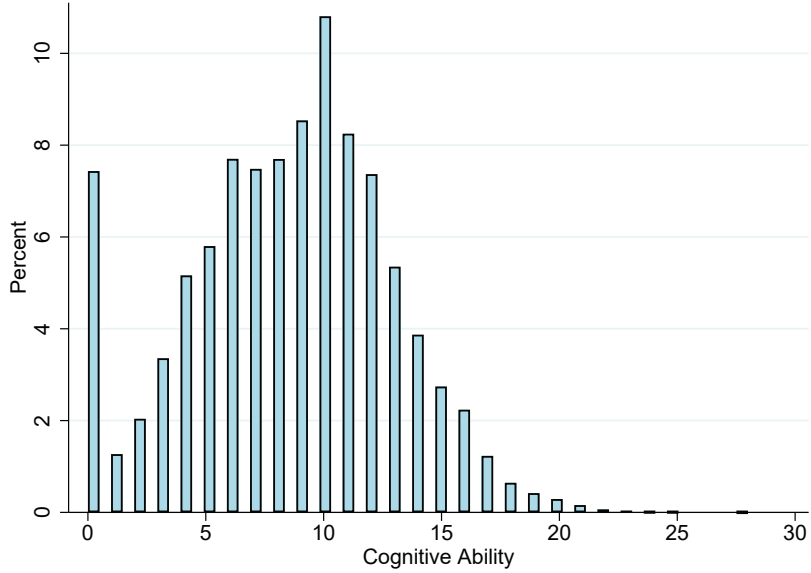
Figure 10: Cognitive ability of the child (SDT test 30)

the interview.

Going beyond the cognitive characteristics we included information about the locus of control complemented with several personality traits corresponding to the Five Factor Model (FFM) (McCrae and Costa (1999)). Information about the locus of control refers to the individual's perception of the relation between effort, luck and inborn abilities. The basic five personality traits include: openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism. Several studies in the literature find that some of the non-cognitive personality traits are relevant for economic success and labour market outcomes (see Anger and Heineck (2010)). Moreover, if there is a strong correlation in these variables between children, parents and their friends than the coefficient of H_{it} can be biased. In order to account for the locus of control we used an answer to the question "In comparison with others, I haven't achieved what I deserved to achieve" provided in SOEP waves 2005, 2010, 2015. The answers were given on a scale 1-7, so we normalized them to the range [0..1] to construct variable "External Locus of Control I". The average of this variable is 0.463 and it is quite dispersed with a substantial variation in the answers. In addition, we used the answers to the question "The abilities we have are more important than the efforts we make" to construct a normalized variable "External Locus of Control II". The answers to this question are more concentrated around the mean equal to 0.689, so there is less variation. We merged both variables to the individual observations in years 2006, 2010 and 2016 by using the personal ID numbers of the respondents. Information about the locus of control is included in the main regression specifications, however, we find a minor role of the big five personality traits for the schooling decision, so the corresponding regressions with these variables are delegated to the appendix. Our findings coming from the main regression specification remain robust to adding information on personality traits.

In order to understand the impact of variable H_{it} on the probability of obtaining higher education for the child in family i we estimate the following logistic regression:

$$\eta_{it}^* = \beta_0 + \beta_1 H_{it} + \beta_2 H_{it}^2 + \beta_3 S_{it}^M + \beta_4 S_{it}^F + Z_{it}'\delta + X_{it}'\gamma + \tau_t + \varepsilon_{it} \tag{2}$$

where η_{it}^* is the latent index, such that the probability of obtaining higher education for the child $P\{\eta = 1\}$ is equal to $P\{\eta^* > 0\}$. H_{it} is a share of high skill friends of family i. S_{it}^M and S_{it}^F are the years of schooling of the mother and the father respectively. Variables Z_{it}' include a set of cognitive (SDT test) and non-cognitive characteristics (locus of control), τ_t are the time fixed effects. X_{it} is a vector of control variables including age, age squared, gender, immigration background and the presence of siblings. Based on the the theoretical model we expect $\beta_1 > 0$, while $\beta_2 \leq 0$ implying that the generation of children gains from a higher share of high skill connections in the family. However, the marginal gain should be declining according to the model and may even turn negative when the number of high skill friends becomes large meaning that there might exist a maximum in the probability of obtaining higher education with respect to H_{it}. Variables S_{it}^M and S_{it}^F are expected to have a positive impact since better educated parents can provide better support to their children in terms of income and information.

One special feature of our data is that SOEP is a household survey, thus, all children in a family receive a separate questionnaire when they become adults. This happens even if they move out of the parental household. Therefore, by construction some individuals in the sample are siblings. In order to control for this we use cluster-robust standard errors on a household level meaning the original household of the parents.

Our results are presented in table 8. In column (1) we predict the probability of obtaining higher education by controlling for the education of parents and the number of high skill friends for each of the two parents separately. As expected all of these variables are positive and significant. In column (2) we use a combined number of high skill friends H_{it} and in column (3) we add a squared term H_{it}^2. The evidence supports the theoretical expectation that there is a positive association between the number of high skill parental friends and the educational attainment of the child ($\beta_1 > 0$) but the marginal gain is declining ($\beta_2 < 0$). We find that the pivot point for the the marginal gain to be negative ($\beta_1/2\beta_2$) is above 1, so the marginal gain remains positive in the relevant range $H_{it} \in [0..1]$.

In columns (4) and (5) we control for cognitive abilities of the child and both parents. Our proxy for own cognitive abilities is marginally significant in column (5) implying that individuals performing better in the symbol digit test (SDT) are more likely to have higher education. Yet, we find that the contribution of parental cognitive abilities is not significant. This result should be taken with care since the number of observations in columns (4) and (5) is rather small. Further, in columns (6) and (7) we include variables corresponding to the locus of control. The first one has a significant negative impact, meaning that individuals who believe that they didn't achieve as much as others are those with lower educational attainment. This is intuitive and shows that some people indirectly regret not having studied longer. The coefficient of the second variable is negative but only marginally significant in regression (6).

Concluding this section, we find that the share of high skill friends in the networks of parents is positively associated with a probability of obtaining higher education for the child, though the marginal impact of this variable is declining. This result is robust to controlling for cognitive abilities of the child and parents and some non-cognitive characteristics. In addition to the logistic regression we also used an OLS specification predicting the years of schooling of the child. The corresponding results are presented in the appendix and show that our findings are robust to the alternative model specification. Overall, our empirical findings are in line with the

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
edu_H							
Share of HS Friends, Mother	1.220*** (0.178)						
Share of HS Friends, Father	1.300*** (0.182)						
Years of Schooling Mother (0-13)	0.138** (0.054)	0.137** (0.054)	0.144*** (0.054)	0.109 (0.103)	0.138 (0.124)	0.178 (0.119)	0.159 (0.103)
Years of Schooling Father (0-13)	0.209*** (0.053)	0.210*** (0.053)	0.207*** (0.053)	0.196** (0.098)	0.150 (0.109)	0.152 (0.130)	0.187 (0.116)
Share of HS Friends, parents		2.521*** (0.226)	3.384*** (0.351)	3.067*** (0.773)	2.965*** (0.892)	2.757*** (0.957)	2.886*** (0.813)
Squared Share of HS Friends, parents			-1.171*** (0.379)	-1.417* (0.809)	-1.684* (0.964)	-1.875* (1.063)	-1.405* (0.843)
Cognitive Ability				0.035 (0.023)	0.043* (0.026)	0.055* (0.028)	0.043* (0.025)
Cognitive Ability Mother					0.011 (0.034)	0.007 (0.037)	
Cognitive Ability Father					-0.009 (0.035)	0.003 (0.038)	
Ext. Locus of Control I						-1.392** (0.551)	-1.060** (0.483)
Ext. Locus of Control II						-1.121* (0.592)	-0.862 (0.545)
Constant	-7.999*** (0.946)	-8.003*** (0.946)	-8.136*** (0.959)	-8.850*** (1.876)	-8.275*** (2.168)	-6.157*** (2.238)	-7.117*** (2.018)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Log-likelihood	-1672.60	-1672.66	-1666.27	-355.68	-282.61	-238.52	-307.66
Pseudo-R squared	0.23	0.23	0.23	0.19	0.17	0.19	0.21
N. of observations	3186	3186	3186	671	538	462	588

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Notes: Logistic regression with higher education of the child as dependent variable. Control variables include gender, age, age squared, first generation immigrant background, second generation immigrant background, the number of siblings and time fixed effects for years 2011 and 2016. Year 2006 is reference category. Cluster-robust standard errors on a household level.

Table 8: Logistic regression predicting higher education of the child

main prediction of the model developed above, in particular proposition 4 and figure 5.

As a form of discussion, we note that the positive link between parental friends and educational attainment of the child could have different reasons and explanations. Whereas, in our theoretical model it is the advantage of having more precise information (via the network) which contributes to more schooling, other studies mention alternative explanations. For example, parental contacts could help children of their friends by sending information about open vacancies at their employers and/or providing a job referral (see Kramarz and Scans (2014)). Thus, children in high skill families may decide to study longer since they expect this kind of support from the network of parental connections. Which of these alternative explanations is more relevant remains a question for future research.

6 Conclusion

In this paper we develop a theoretical and empirical framework to study the role of parental networks for the educational decisions of children and intergenerational mobility. Children observe the skill types and get noisy signals of wages from the social contacts of their parents. They use this information to form expectations about the market skill premium, which is otherwise unobservable. The schooling is modeled as a binary variable, hence children decide to become either high or low skilled based on their estimated return to schooling, inborn ability and the cost of education. We find that the skill premium estimated by children in low skill families

is more dispersed, indicating higher uncertainty, compared to the estimates of children in high skill families. This happens if social networks exhibit skill homophily and high skill parents are a minority in the economy. As a result, children in low skill families are less likely to obtain higher education compared to children in high skill families giving rise to a positive intergenerational schooling correlation. The correlation is amplified if high skill wages are more dispersed than low skill wages. We view this mechanism as complementary to the traditional explanations described in the literature on intergenerational mobility.

Further, the model addresses a conceptual issue that wages of parents and close family members and friends (strong ties) are typically observed with high precision, whereas wages of former colleagues and acquaintances (weak ties) are observed with lower precision. In this case expectation formation of the young generation suffers from the problem of heteroscedasticity. We prove that the optimal response of the youth to cope with this problem is to assign a higher weight to the precise observations of parental wages and lower weights to the remaining observations (weighted least squares), giving rise to the efficient estimates of the return to schooling.

We calibrate the model to match information about wages, social networks and schooling attainment observed in the German data (SOEP). Depending on the cost of education the new mechanism described by our model can contribute up to 15% of the observed intergenerational correlation coefficient. Based on the calibration we find that the correlation coefficient has low sensitivity to the degree of uncertainty about the return to schooling and the size of the social network, but it is very sensitive to the share of high skill parents in the economy and the degree of skill homophily. The data reveals that there is strong homophily by skill in the German society but it was stable over the last two decades, thus we direct our attention to the link between the share of high skill parents/families and the intergenerational correlation coefficient and find that the inverse U-shape relationship between these two variables predicted by the model is confirmed by the empirical data.

Finally, we test the main prediction of the model that the probability of obtaining higher education for the child should be increasing in the share of high skill parental friends with a declining marginal gain. For high skill families the marginal gain may even become negative but numerically the decline in the probability is small. We test these predictions by using data from the German Socio-Economic Panel (SOEP). In particular, we predict the probability of obtaining higher education for the child controlling for the education of parents, demographics of the child as well as a set of cognitive and non-cognitive characteristics. We find that the share of high skill parental friends is positive and significant with a negative quadratic term reflecting a falling marginal gain. This finding is in line with the main theoretical prediction of the model. In addition, we find that cognitive abilities of the child and the locus of control have moderate effects for predicting educational attainment. However, cognitive abilities of parents and the big five personality traits of the child do not seem to play an important role.

7   Appendix

Proof of Proposition 1.

$$\frac{V[w_i^H]}{n^{LH}} + V[w_i^L]\Big(s^2 + \frac{(1-s)^2}{n^{LL}}\Big) \geq V[w_i^H]\Big(s^2 + \frac{(1-s)^2}{n^{HH}}\Big) + \frac{V[w_i^L]}{n^{HL}}$$

which is equivalent to

$$V[w_i^H]\Big(\frac{1}{n^{LH}} - s^2 - \frac{(1-s)^2}{n^{HH}}\Big) \geq V[w_i^L]\Big(\frac{1}{n^{HL}} - s^2 - \frac{(1-s)^2}{n^{LL}}\Big)$$

After substituting the expressions for $n^{LH}, n^{LH}, n^{HL}, n^{HH}$:

$$V[w_i^H]\Big(\frac{1}{\varphi h(1-\alpha)} - s^2 - \frac{(1-s)^2}{\varphi h}\Big) \geq V[w_i^L]\Big(\frac{1}{\varphi(1-h)(1-\alpha)} - s^2 - \frac{(1-s)^2}{\varphi(1-h)}\Big)$$

which is equivalent to:

$$\frac{V[w_i^H]}{h}\Big(1 - (1-s)^2(1-\alpha) - \varphi s^2 h(1-\alpha)\Big) \geq \frac{V[w_i^L]}{(1-h)}\Big(1 - (1-s)^2(1-\alpha) - \varphi s^2(1-h)(1-\alpha)\Big)$$

Thus we have to compare the ratio of weighted variances with the following expression:

$$\frac{V[w_i^H]}{V[w_i^L]}\frac{(1-h)}{h} \quad \text{vs.} \quad 1 - \frac{\varphi(1-\alpha)(1-2h)s^2}{1 - (1-s)^2(1-\alpha) - \varphi s^2 h(1-\alpha)} \equiv G(s)$$

The denominator of $G(s)$ can be rewritten as:

$$-(1-\alpha)(\varphi h + 1)s^2 + 2(1-\alpha)s + \alpha$$

which is a quadratic equation in s with one negative root and one positive, the latter is:

$$s* = \frac{1}{\varphi h + 1}\Big(1 + \sqrt{1 + \frac{\alpha(\varphi h + 1)}{1 - \alpha}}\Big)$$

Thus, function $G(s)$ has a vertical asymptote at s^*, which can be larger or smaller than 1. Next differentiate $G(s)$ with respect to s:

$$\frac{G'(s)}{\varphi(1-\alpha)(1-2h)} = -\frac{2s[1 - (1-s)^2(1-\alpha) - \varphi s^2 h(1-\alpha)] - s^2[2(1-s)(1-\alpha) - 2\varphi s h(1-\alpha)]}{(1 - (1-s)^2(1-\alpha) - \varphi s^2 h(1-\alpha))^2}$$

$$= -\frac{2s(1 - (1-s)(1-\alpha))}{(1 - (1-s)^2(1-\alpha) - \varphi s^2 h(1-\alpha))^2} < 0$$

This means the following. For $s < s^*$, function $G(s)$ starts at a value $G(0) = 1$, is decreasing for higher values of s and converges to $-\infty$ when s is converging to s^* from below. In this range we know that $G(s) \leq 1$, hence any value of the weighted variance ratio $(V[w_i^H]/V[w_i^L])((1-h)/h)$ above 1 would lead to the fact that the variance of the skill premium Δ_j^L is greater than that of Δ_j^H. This is the first sufficient condition. It is the only condition if $s^* > 1$.

Next consider $s^* < s < 1$, then function $G(s)$ is decreasing from ∞ down to $G(1) > 1$, thus the relevant sufficient condition becomes:

$$\frac{V[w_i^H]}{V[w_i^L]}\frac{(1-h)}{h} < G(1) = 1 + \frac{\varphi(1-\alpha)(1-2h)}{\varphi h(1-\alpha) - 1}$$

This is a second sufficient condition (by assumption $h < 0.5$ and $n^{LH} = \varphi h(1-\alpha) > 1$). $\square$

Proof of Lemma 2. The first order derivatives of $V[\Delta^L]$ and $V[\Delta^H]$ with respect to φ become:

$$\frac{\partial V[\Delta^L]}{\partial \varphi} = -\frac{V[w_i^H]}{\varphi^2 h(1-\alpha)} - \frac{V[w_i^L](1-s)^2}{\varphi^2(1-h)} = -\frac{1}{\varphi^2}\Big(\frac{V[w_i^H]}{h(1-\alpha)} + \frac{V[w_i^L](1-s)^2}{(1-h)}\Big) < 0$$

$$\frac{\partial V[\Delta^H]}{\partial \varphi} = -\frac{V[w_i^H](1-s)^2}{\varphi^2 h} - \frac{V[w_i^L]}{\varphi^2(1-h)(1-\alpha)} = -\frac{1}{\varphi^2}\Big(\frac{V[w_i^H](1-s)^2}{h} + \frac{V[w_i^L]}{(1-h)(1-\alpha)}\Big) < 0$$

One can also observe that

$$\frac{V[w_i^H]}{h(1-\alpha)} + \frac{V[w_i^L](1-s)^2}{(1-h)} > \frac{V[w_i^H](1-s)^2}{h} + \frac{V[w_i^L]}{(1-h)(1-\alpha)}$$

which is equivalent to

$$\left(\frac{1}{(1-\alpha)} - (1-s)^2\right)\left(\frac{V[w_i^H]}{h} - \frac{V[w_i^L]}{(1-h)}\right) > 0 \quad \Rightarrow \quad \frac{\partial V[\Delta^L]}{\partial\varphi} < \frac{\partial V[\Delta^H]}{\partial\varphi}.$$

Next differentiating $V[\Delta^L]$ and $V[\Delta^H]$ with respect to α become:

$$\frac{\partial V[\Delta^L]}{\partial\alpha} = V[w_i^H]\frac{\varphi h}{(n^{LH})^2} = \frac{V[w_i^H]}{\varphi h(1-\alpha)^2} > 0$$

$$\frac{\partial V[\Delta^H]}{\partial\alpha} = V[w_i^L]\frac{\varphi(1-h)}{(n^{HL})^2} = \frac{V[w_i^L]}{\varphi(1-h)(1-\alpha)^2} > 0 \quad \Rightarrow \quad \frac{\partial V[\Delta^L]}{\partial\alpha} > \frac{\partial V[\Delta^H]}{\partial\alpha}.$$

Further, differentiating $\Phi^k(c)$ $(k = L, H)$ with respect to φ and α, we get:

$$\frac{\partial\Phi^i(c)}{\partial\varphi} = \frac{1}{2\sigma^i}\frac{\partial\Phi^i(c)}{\partial\sigma^i}\frac{\partial V[\Delta^i]}{\partial\varphi} < 0 \quad \text{and} \quad \frac{\partial\Phi^i(c)}{\partial\alpha} = \frac{1}{2\sigma^i}\frac{\partial\Phi^i(c)}{\partial\sigma^i}\frac{\partial V[\Delta^i]}{\partial\alpha} > 0.$$

This means that in both types of families posses more precise information about skill premium and more children obtain education when the social network of parents exhibit less degree of homophily or the average size of the network increases. Otherwise, the education gap in both types of families is increasing when the social network exhibits strong degree of homophily or the average size of network is small. □

Proof of Proposition 2. Note that $E[\xi] = h$ and $E[\eta] = \tilde{h}$. Given Table 1, we can compute covariance of ξ and η using the fact that $cov[\xi,\eta] = E[\xi\eta] - E[\xi]E[\eta]$ where

$$\begin{aligned}
E[\xi\eta] &= \sum_{x=\{0,1\}}\sum_{y=\{0,1\}} xyP(\xi = x, \eta = y)\\
&= 0\times 0\times P(\xi=0,\eta=0) + 1\times 0\times P(\xi=1,\eta=0)\\
&+ 0\times 1\times P(\xi=0,\eta=1) + 1\times 1\times P(\xi=1,\eta=1) = hp(1-\Phi^H(c)).
\end{aligned}$$

Then, $cov[\xi,\eta] = hp(1-\Phi^H(c)) - h\tilde{h}$. Substituting in $\tilde{h}$, we obtain

$$\begin{aligned}
cov[\xi,\eta] &= h(p(1-\Phi^H(c)) - p(1-\Phi^H(c)) + p(1-h)\Delta\Phi(c)) = (1-h)hp\Delta\Phi(c)\\
V[\xi] &= (1-h)^2h + (0-h)^2(1-h) = (1-h)h(1-h+h) = (1-h)h\\
V[\eta] &= (1-\tilde{h})^2\tilde{h} + (0-\tilde{h})^2(1-\tilde{h}) = (1-\tilde{h})\tilde{h}(1-\tilde{h}+\tilde{h}) = (1-\tilde{h})\tilde{h}\\
\rho &= \frac{(1-h)hp\Delta\Phi(c)}{\sqrt{h(1-h)\tilde{h}(1-\tilde{h})}} = p\Delta\Phi(c)\sqrt{\frac{(1-h)h}{(1-\tilde{h})\tilde{h}}}
\end{aligned}$$

□

Proof of Proposition 3. In order to find a variance minimizing point s^k, $k = L, H$ we differentiate $V[\Delta_j^k]$ with respect to s:

$$\begin{aligned}
2sV[w_i^k] - \frac{2(1-s)}{n^{kk}}V[w_i^k] - \frac{2(1-s)\sigma_{\epsilon k}^2}{n^{kk}} &= 0\\
s(V[w_i^k]n^{kk} + V[w_i^k] + \sigma_{\epsilon k}^2) - V[w_i^k] - \sigma_{\epsilon k}^2 &= 0 \quad \Rightarrow \quad s^k = \frac{V[w_i^k] + \sigma_{\epsilon k}^2}{n^{kk}V[w_i^k] + V[w_i^k] + \sigma_{\epsilon k}^2}
\end{aligned}$$

which gives rise to values s^L and s^H respectively. Note that the variance of Δ_j^k is minimized for a value of s which is minimizing the variance of the average wage w_j^{kk}. Next we show that these values of s coincide with the weighted LS estimator. In the following we suppress the skill index k for the ease of exposition and use the following notation: $E[w_j^{kk}] = E[w_i^k] = \mu$ and $V[w_i^k] = \sigma_w^2$. Without loss of generality, let the wage of the parent be observation 0, and wage observations of social contacts be numbered $l = 1..n$, so we get the system of equations:

$$
\begin{aligned}
w_0 &= \mu + \eta_0, \quad \text{where} \quad V[\eta_|] = \sigma_w^2 \\
w_1 &= \mu + \eta_1 + \epsilon_1 \quad \text{where} \quad V[\eta_1 + \epsilon_1] = \sigma_w^2 + \sigma_\epsilon^2 \\
&\vdots \\
w_n &= \mu + \eta_n + \epsilon_n \quad \text{where} \quad V[\eta_n + \epsilon_n] = \sigma_w^2 + \sigma_\epsilon^2
\end{aligned}
$$

These equations show that the wage observation of the parent w_1 does not have the noise component of the error term and is more precise in the sense that $V[\eta_0] < V[\eta_l + \epsilon_l]$, $l = 1..n$ indicating the problem of heteroscedasticity. The best unbisased estimator of μ is then given by the weighted LS estimator $\hat{\mu} = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}Y$, where X is a vector of ones and Y is a vector of wages. The weighting matrix Ω and its inverse are given by:

$$
\Omega = \begin{pmatrix} \sigma_w^2 & 0 & \cdots & 0 \\ 0 & \sigma_w^2 + \sigma_\epsilon^2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \sigma_w^2 + \sigma_\epsilon^2 \end{pmatrix} \quad \rightarrow \quad \Omega^{-1} = \begin{pmatrix} \frac{1}{\sigma_w^2} & 0 & \cdots & 0 \\ 0 & \frac{1}{\sigma_w^2 + \sigma_\epsilon^2} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \frac{1}{\sigma_w^2 + \sigma_\epsilon^2} \end{pmatrix}
$$

The WLS estimator $\hat{\mu}$ is then given by:

$$
\begin{aligned}
\hat{\mu} &= \left(\frac{1}{\sigma_w^2} + \sum_{l=1}^n \frac{1}{\sigma_w^2 + \sigma_\epsilon^2} \right)^{-1} \left(\frac{w_0}{\sigma_w^2} + \sum_{l=1}^n \frac{w_l}{\sigma_w^2 + \sigma_\epsilon^2} \right) \\
&= \frac{\sigma_w^2(\sigma_w^2 + \sigma_\epsilon^2)}{\sigma_w^2 + \sigma_\epsilon^2 + n\sigma_w^2} \left(\frac{w_0(\sigma_w^2 + \sigma_\epsilon^2) + \sigma_w^2 \sum_{l=1}^n w_l}{\sigma_w^2(\sigma_w^2 + \sigma_\epsilon^2)} \right) = \frac{w_0(\sigma_\epsilon^2 + \sigma_w^2)}{n\sigma_w^2 + \sigma_\epsilon^2 + \sigma_w^2} + \frac{n\sigma_w^2(\frac{1}{n}\sum_{l=1}^n w_l)}{n\sigma_w^2 + \sigma_\epsilon^2 + \sigma_w^2}
\end{aligned}
$$

Given our notation ($V[w_i^k] = \sigma_w^2$, $n^{kk} = n$, $\sigma_{\epsilon k}^2 = \sigma_\epsilon^2$) we can see that the weighted LS estimator implies a weight on the wage of the parent equal to s^k. $\quad\square$

Proof of proposition 4.

$$
\begin{aligned}
\frac{\partial V[\Delta_j^L]}{\partial h} &= -\frac{(1+z)V[w_i^H]}{\varphi(1-\alpha)h^2} - \frac{(1+z)\varphi V[w_i^L]}{(n^{LL}+1+z)^2} = (1+z)\left[\frac{\varphi V[w_i^L]}{(\varphi(1-h)+1+z)^2} - \frac{V[w_i^H]}{\varphi(1-\alpha)h^2} \right] \\
&= (1+z)\left[\frac{\varphi^2(1-\alpha)h^2 V[w_i^L] - (\varphi(1-h)+1+z)^2 V[w_i^H]}{\varphi(1-\alpha)h^2(\varphi(1-h)+1+z)^2} \right] < 0
\end{aligned}
$$

where we use the assumptions that: $V[w_i^H]/V[w_i^L] > h/(1-h)$ and $0 < h < 0.5$. This allows us to use the following inequalities:

$$
\frac{V[w_i^H]}{V[w_i^L]} > \frac{h}{(1-h)} > \frac{h^2}{(1-h)^2} > \frac{\varphi^2(1-\alpha)h^2}{(\varphi(1-h)+1+z)^2}.
$$

This means that the share of children obtaining higher education in low skill families is increasing in h, that is $p(1 - \Phi^L(c))$ because $\frac{\partial \Phi^L(c)}{\partial h} = \frac{1}{2\sigma^L} \frac{\partial \Phi^L(c)}{\partial \sigma^L} \frac{\partial V[\Delta_j^L]}{\partial h} < 0$. Next we consider children in high skill

families:

$$\frac{\partial V[\Delta_j^H]}{\partial h} = -\frac{(1+z)\varphi V[w_i^H]}{(n^{HH}+1+z)^2} + \frac{(1+z)V[w_i^L]}{\varphi(1-\alpha)(1-h)^2}$$

$$= (1+z)\Big[\frac{V[w_i^L]}{\varphi(1-\alpha)(1-h)^2} - \frac{\varphi V[w_i^H]}{(\varphi h + 1 + z)^2}\Big]$$

Let us define h^* as follows:

$$\frac{\varphi^2(1-\alpha)(1-h^*)^2}{(\varphi h^* + 1 + z)^2} = \frac{V[w_i^L]}{V[w_i^H]} \quad \Rightarrow \quad h^* = \frac{\varphi\sqrt{(1-\alpha)V[w_i^H]} - (1+z)\sqrt{V[w_i^L]}}{\varphi(\sqrt{(1-\alpha)V[w_i^H]} + \sqrt{V[w_i^L]})}$$

One can see that $h^* < 0.5$ if $(1-\alpha)\varphi^2/(\varphi + 2(1+z))^2 < V[w_i^L]/V[w_i^H]$. If it holds then:

$$\begin{cases} \frac{\partial V[\Delta_j^H]}{\partial h} < 0 \quad \Rightarrow \frac{\partial \Phi^H(c)}{\partial h} < 0 \quad \text{if} \quad h \in (0, h^*) \\ \frac{\partial V[\Delta_j^H]}{\partial h} > 0 \quad \Rightarrow \frac{\partial \Phi^H(c)}{\partial h} > 0 \quad \text{if} \quad h \in (h^*, 0.5) \end{cases}$$

This means that the share of children obtaining education in high skill families $p(1 - \Phi^H(c))$ is decreasing in h for $h \in [h^* ..0.5]$ if $(1-\alpha)\varphi^2/(\varphi + 2(1+z))^2 < V[w_i^L]/V[w_i^H]$. Therefore, the mobility ratio m is decreasing in h in this range. The effect on the mobility ratio is ambiguous if $h < h^*$. $\quad\square$

Extension: Ability transmission In this section we extend the model by incorporating a positive correlation in the abilities of parents and children. The goal of developing this extension is to analyze the interaction between the pure ability driven correlation in schooling and the network driven correlation described above. In addition, in this section we investigate the fit of an overall model with both components to the empirical data.

Let variable ζ denote a binary indicator for the ability of the child, such that $\zeta = 1$ for high ability children and $\zeta = 0$ for low ability children. Further, $p^H > p$ denotes the probability that a child born in a high skill family has high ability and $p^L < p$ be the probability that a child born in low skill family has high ability. The fraction of high ability children is then $E[\zeta] = hp^H + (1-h)p^L$ and the covariance between the skill level of the parent ξ and the ability of the child ζ is given by:

$$cov[\xi, \zeta] = E[\xi, \zeta] - E[\xi]E[\zeta] = hp^H - h(hp^H + (1-h)p^L) = h(1-h)(p^H - p^L)$$

Thus the covariance (and the correlation coefficient) between ξ and ζ are strongly driven by the difference $p^H - p^L$. Following the literature we attribute this correlation to the combined effect of nature and nurture, meaning that this correlation can be driven by the genetic transmission of abilities (nature), but can also stem from the more efficient help of high skill parents to their children (nurture).

Next consider the intergenerational covariance in schooling $cov(\xi, \eta)$ in the extended model with ability and network components. The number of children acquiring education in high skill families is given by $p^H(1 - \Phi^H(c))$ and the number of children in low skill families acquiring education is given by $p^L(1 - \Phi^L(c))$. So the matrix of transition probabilities becomes:

	$\eta = 1$	$\eta = 0$	Σ
$\xi = 1$	$hp^H(1 - \Phi^H(c))$	$h(1 - p^H(1 - \Phi^H(c)))$	h
$\xi = 0$	$(1-h)p^L(1 - \Phi^L(c)))$	$(1-h)(1 - p^L(1 - \Phi^L(c)))$	$1 - h$
	$\bar{h}$	$1 - \bar{h}$	1

Thus, the number of high skill children in the combined model is given by $\bar{h} = hp^H(1 - \Phi^H(c)) + (1 -$

$h)p^L(1 - \Phi^L(c)))$. Given this notation we can derive the intergenerational covariance in schooling in the combined model in the following way:

$$
\begin{aligned}
cov(\xi, \eta) &= E[\xi\eta] - E[\xi]E[\eta] = hp^H(1 - \Phi^H(c)) - h\bar{h} \\
&= h(1 - h)[p^H(1 - \Phi^H(c)) - p^L(1 - \Phi^L(c))]
\end{aligned}
$$

We can see that it is increasing in the number of children acquiring education in high skill families $p^H(1 - \Phi^H(c))$ and decreasing in the number of children acquiring education in low skill families $p^L(1 - \Phi^L(c))$. Next we decompose the schooling covariance $cov(\xi, \eta)$ into the ability component $cov[\xi, \zeta]$, the network component and their interaction term:

$$
\begin{aligned}
cov(\xi, \eta) &= h(1 - h)[p^H(1 - \Phi^H(c)) - p^L(1 - \Phi^L(c))] \\
&= \underbrace{h(1 - h)(p^H - p^L)}_{\text{ability effect}} + \underbrace{h(1 - h)p\Delta\Phi(c)}_{\text{network effect}} + \underbrace{h(1 - h)[(p^L - p)\Phi^L(c) - (p^H - p)\Phi^H(c)]}_{\text{interaction effect} <0}
\end{aligned}
$$

This equation shows that the intergenerational schooling covariance can be decomposed into (1) the ability component driven by the difference $p^H - p^L$, (2) the network component driven by the difference $\Phi^L(c) - \Phi^H(c)$ and (3) the interaction term. We can see that the interaction term is always negative in the presence of a positive ability correlation because $p^L < p$ and $p^H > p$. This means that stronger ability correlation crowds out the network effect. The intuitive explanation for this finding is the following: The network effect is largely driven by high ability children born in low skill families. These children are facing strong uncertainty about the market for high skill workers and are likely to make a mistake by not obtaining the education. However, as the ability correlation gets stronger, the number of high ability children born in low skill families is decreasing, so the number of children, who can potentially make a mistake by not studying, is also decreasing. This explains the negative sign of the interaction effect and the crowding out effect.

I see myself as someone who	Personality trait
is original	Openness to Experience (50%)
values artistic experiences	Openness to Experience (50%)
is a thorough worker	Conscientiousness (50%)
carries out tasks efficiently	Conscientiousness (50%)
is sociable	Extraversion (50%)
is communicative	Extraversion (50%)
is able to forgive	Agreeableness (50%)
is friendly with others	Agreeableness (50%)
worries a lot	Neuroticism (50%)
is somewhat nervous	Neuroticism (50%)
Compared to other people, I have not achieved what I deserve	External locus of control I
What a person achieves in life is above all a question of fate or luck	External locus of Control II

Table 9: SOEP Questions on Personality Traits and the Locus of Control

	Mean	SD	Min	Max	N
Conscientiousness	0.857	0.127	0	1	34349
Extraversion	0.770	0.176	0	1	34425
Agreeableness	0.812	0.140	0	1	34398
Openness	0.646	0.189	0	1	34153
Neuroticism	0.583	0.202	0	1	34397

Table 10: Personality traits: descriptive statistics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
edu_H							
Share of HS Friends, Mother	1.220***						
	(0.178)						
Share of HS Friends, Father	1.300***						
	(0.182)						
Years of Schooling Mother (0-13)	0.138**	0.137**	0.144***	0.109	0.138	0.178	0.159
	(0.054)	(0.054)	(0.054)	(0.103)	(0.124)	(0.119)	(0.103)
Years of Schooling Father (0-13)	0.209***	0.210***	0.207***	0.196**	0.150	0.152	0.187
	(0.053)	(0.053)	(0.053)	(0.098)	(0.109)	(0.130)	(0.116)
Survey Year=2011	0.171**	0.172**	0.174**	-0.080	-0.124	-0.143	-0.100
	(0.085)	(0.085)	(0.085)	(0.237)	(0.266)	(0.309)	(0.266)
Survey Year=2016	0.324***	0.323***	0.311***	0.017	0.099	-0.026	-0.094
	(0.103)	(0.103)	(0.103)	(0.239)	(0.272)	(0.299)	(0.259)
Male gender	-0.242**	-0.243**	-0.249**	-0.535**	-0.568**	-0.514*	-0.459*
	(0.111)	(0.111)	(0.111)	(0.223)	(0.250)	(0.284)	(0.246)
Siblings	-0.484***	-0.483***	-0.472***	-0.409	-0.497	-0.516	-0.470
	(0.179)	(0.179)	(0.181)	(0.388)	(0.418)	(0.463)	(0.430)
Age	0.229***	0.230***	0.226***	0.281***	0.252**	0.172	0.217**
	(0.044)	(0.044)	(0.044)	(0.100)	(0.118)	(0.118)	(0.104)
Age Squared	-0.003***	-0.003***	-0.003***	-0.004**	-0.003*	-0.002	-0.003*
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
First Generation Immigrant	1.023**	1.020**	1.032**	1.072	0.302	0.788	1.350*
	(0.447)	(0.446)	(0.438)	(0.764)	(1.035)	(1.196)	(0.805)
Second Generation Immigrant	0.186	0.187	0.225	0.494	0.487	-0.041	-0.115
	(0.270)	(0.270)	(0.272)	(0.693)	(0.772)	(1.015)	(0.886)
Share of HS Friends, parents		2.521***	3.384***	3.067***	2.965***	2.757***	2.886***
		(0.226)	(0.351)	(0.773)	(0.892)	(0.957)	(0.813)
Squared Share of HS Friends, parents			-1.171***	-1.417*	-1.684*	-1.875*	-1.405*
			(0.379)	(0.809)	(0.964)	(1.063)	(0.843)
Cognitive Ability				0.035	0.043*	0.055*	0.043*
				(0.023)	(0.026)	(0.028)	(0.025)
Cognitive Ability Mother					0.011	0.007	
					(0.034)	(0.037)	
Cognitive Ability Father					-0.009	0.003	
					(0.035)	(0.038)	
Ext. Locus of Control I						-1.392**	-1.060**
						(0.551)	(0.483)
Ext. Locus of Control II						-1.121*	-0.862
						(0.592)	(0.545)
Constant	-7.999***	-8.003***	-8.136***	-8.850***	-8.275***	-6.157***	-7.117***
	(0.946)	(0.946)	(0.959)	(1.876)	(2.168)	(2.238)	(2.018)
Log-likelihood	-1672.60	-1672.66	-1666.27	-355.68	-282.61	-238.52	-307.66
Pseudo-R squared	0.23	0.23	0.23	0.19	0.17	0.19	0.21
N. of observations	3186	3186	3186	671	538	462	588

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Year 2006 is reference category. Cluster-robust standard errors on a household level.

Table 11: Logistic regression predicting higher education of the child (full size)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
main							
Share of HS Friends, parents	2.567***	2.268**	3.191***	2.734***	2.797***	2.571***	2.413***
	(0.957)	(0.992)	(0.315)	(0.680)	(0.875)	(0.695)	(0.823)
Squared Share of HS Friends, parents	-0.118	-0.100	-1.570***	-1.431**	-0.800	-1.436**	-0.768
	(1.028)	(1.048)	(0.288)	(0.621)	(0.761)	(0.653)	(0.785)
Years of Schooling Mother (0-13)	0.209**	0.247**	0.124***	0.225**	0.273**	0.224**	0.256**
	(0.102)	(0.109)	(0.041)	(0.105)	(0.123)	(0.106)	(0.116)
Years of Schooling Father (0-13)	0.253*	0.247*	0.096***	0.091	0.073	0.066	0.043
	(0.134)	(0.138)	(0.029)	(0.068)	(0.071)	(0.066)	(0.069)
Male gender	-0.545*	-0.420	-0.172*	-0.505***	-0.547**	-0.513***	-0.513**
	(0.295)	(0.310)	(0.095)	(0.177)	(0.260)	(0.191)	(0.257)
Cognitive Ability	0.065**	0.055*		0.055***	0.076***	0.063**	0.075***
	(0.029)	(0.029)		(0.021)	(0.027)	(0.025)	(0.028)
Conscientiousness	-1.481	-1.235			0.903		1.657
	(1.283)	(1.364)			(1.620)		(1.747)
Extraversion	0.352	0.348			-0.390		-0.445
	(0.912)	(0.907)			(0.679)		(0.699)
Agreeableness	-1.132	-1.054			-0.272		-0.412
	(1.093)	(1.125)			(1.179)		(1.183)
Openness	0.795	0.817			-0.433		-0.286
	(0.860)	(0.884)			(0.721)		(0.799)
Neuroticism	-0.632	-0.091			-0.607		-0.142
	(0.741)	(0.761)			(0.619)		(0.524)
Survey Year=2016	-0.234	-0.255	0.383***	0.139	0.038	0.162	0.051
	(0.288)	(0.296)	(0.103)	(0.181)	(0.214)	(0.206)	(0.216)
Siblings	-0.778*	-0.723	-0.437***	-0.649***	-0.622**	-0.643**	-0.481
	(0.415)	(0.443)	(0.120)	(0.247)	(0.316)	(0.275)	(0.342)
Age	0.190*	0.168	0.172***	0.156*	0.064	0.122	0.066
	(0.107)	(0.107)	(0.041)	(0.085)	(0.111)	(0.085)	(0.103)
Age Squared	-0.002	-0.002	-0.003***	-0.002	-0.001	-0.002	-0.001
	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)
First Generation Immigrant	2.126***	2.099***	0.464	0.878**	1.116**	1.131***	1.356***
	(0.813)	(0.784)	(0.302)	(0.415)	(0.542)	(0.366)	(0.479)
Second Generation Immigrant	0.470	0.387	-0.290	-0.455	-0.469	-0.541	-0.585
	(0.912)	(1.056)	(0.227)	(0.449)	(0.461)	(0.499)	(0.522)
Ext. Locus of Control I		-1.919***				-0.765**	-1.540***
		(0.644)				(0.389)	(0.434)
Ext. Locus of Control II		-0.345				-1.559***	-1.551*
		(0.712)				(0.473)	(0.797)
Survey Year=2011			0.349***	0.215		0.192	
			(0.081)	(0.183)		(0.213)	
Constant	-7.469***	-6.783**	5.932***	4.924***	6.168***	7.083***	7.602***
	(2.537)	(2.666)	(0.761)	(1.527)	(2.088)	(1.691)	(2.183)
Log-likelihood	-202.94	-193.63	-5605.65	-1225.69	-775.58	-1068.37	-743.33
Pseudo-R squared	0.26	0.28					
Adjusted R squared			0.35	0.37	0.41	0.41	0.44
N. of observations	417	406	3161	665	413	582	402

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Notes: Columns (1)-(2) correspond to a Logistic regression with higher education of the child as dependent variable.
Columns (3)-(7) correspond to an OLS regression for the child's years of schooling. Cluster-robust standard errors on a household level.

Table 12: Logistic and OLS regression outputs

8 References

ALTONJI G.J., R.K. MANSFIELD (2018) "Estimating Group Effects Using Averages of Observables to Control for Sorting on Unobservables: School and Neighborhood Effects", *American Economic Review*, 108(10): 2902-2946.

ANGER S., HEINECK G. (2010) "The returns to cognitive abilities and personality traits in Germany", *Labour Economics*, 17(3): 535-546.

BECKER G. S., N. THOMES (1986) "Human Capital and the Rise and Fall of Families", *Journal of Labour Economics*, 4(3): 1-39.

BELLEY P., M. FRENETTE, L. LOCHNER (2011) "Post-Secondary Attendance by Parental Income in the US and Canada: What Role for Financial Aid Policy?", *NBER Working Paper Series*, 17218.

BJÖRKLUND A., L. LINDAHL, M.J. LINDQUIST (2010) "What More Than Parental Income, Education and Occupation? An Exploration of What Swedish Siblings of Get from Their Parents", *The B.E. Journal of Economic Analysis & Policy*, 10(1).

BJÖRKLUND A., M. JÄNTTI (1997) "Intergenerational Income Mobility in Sweden Compared to the United States", *American Economic Review*, 87(5): 1009-1018.

BOWLES S., H. GINTIS (2002) "The Inheritance of Inequality", *Journal of Economic Perspective*, 16(3): 2-30.

BLACK S.E, P.J. DEVEREUX, K.G. SALVANES (2009) "Like Father, Like Son? A Note on the Intergenerational Transmission of IQ Scores", *Economics Letters*, 105(1): 138-140.

BRATBERG E., J. DAVIS, B. MAZUMDER, M. NYBOM, D.D. SCHNITZLEIN, K. VAAGE (2016) "A Comparsion of Intergenerational Mobility Curves in Germany, Norway, Sweden, and the US", *The Scandinavian Journal of Economics*, 00(0): 1-30.

BURDETT K., D.T. MORTENSEN (1998) "Wage Differentials, Employer Size and Unemployment", *International Economic Review*, 39(2): 257-273.

CAHUC P., F. FONTAINE (2009) "On the Efficiency of Job Search with Social Networks", *Journal of Public Economic Theory*, 11(3): 411-439.

CARD D. (2001) "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems", *Econometrica*, 65(5): 1127-1160.

CARNEIRO P., J.T. HECKMAN (2002) "The Evidence on Credit Constraints in Post-Secondary Schooling", *The Economic Journal*, 112(482): 705-734.

CALVÓ-ARMENGOL A., M. JACKSON (2009) "Like Father, Like Son: Social Network Externalities and Parent-Child Correlation in Behavior", *American Economic Journal: Microeconomics*, 1(1): 124-150.

CALVÓ-ARMENGOL A., Y. ZENOU (2005) "Job Matching, Social Network and Word-of-Mouth Communication", *Journal of Urban Economics*, 57(3): 500-522.

CHECCI D., A. ICHINO, A. RUSTICHINI (1999) "More Equal But Less Mobile?: Education Financing and Intergenerational Mobility in Italy", *Journal of Public Economics*, 74(3): 351-393.

CHECCI D., L. FLABBI (2007) "Intergenerational Mobility and Schooling Decisions in Germany and Italy: The Impact of Secondary School Tracks", *IZA Discussion Paper*, No. 2876.

CHETTY R., N. HENDREN, P. KLINE, E. SAEZ (2014) "Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States", *The Quarterly Journal of Economics*, 129(4): 1553-1623.

CHETTY R., N. HENDREN, L.F. KATZ (2016) "The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiement", *American Economic Review*, 106(4): 855-902.

CHEVALIER A., K. DENNY, D. MCMAHON (2003) "A Multi-Country Study of Intergenerational Educational Mobility", *Centre for Economic Research Working Paper Series*, 14.

CINGANO F., A. ROSOLIA (2012) "People I Know: Job Search and Social Networks", *Journal of Labor Economics*, 30(2): 291-332.

CORAK M., P. PIRAINO (2011) "The Intergenerational Transmission of Employers", *Journal of Labor Economics*, 29: 37-68.

COOKE L.P. (2003) "A Comparison of Initial and Early Life Course Earnings of the German Secondary Education and Training System", *Economics of Education Review*, 22(1): 79-88.

COUCH K.A, T.A. DUNN (1997) "Intergenerational Correlations in Labor Market Status: A Comparison of the United States and Germany", *Journal of Human Resources*, 32(1): 210-232.

DIAMOND P. A. and J. E. STIGLITZ (1973) "Increase in Risk and in Risk Aversion", MIT *Working paper of the department of economics*, Nr. 97.

DEL BELLO L.C., E. PATACCHINI, Y. ZENOU (2015) "Neighborhood Effects in Education", *IZA Working Paper*, 8956.

DOEPKE M., M. TERTILT (2016) "Families in Macroeconomics", Chapter 23 of *Handbook of Macroeconomics*, Volume 1. North Holland.

EK S., B. HOLMLUND (2010) "Family Job Search, Wage Bargaining, and Optimal Unemployment Insurance", *The B.E. Journal of Economic Analysis and Policy*, 10(1):47.

FLABBI L., J. MABLI (2018) "Household Search or Individual Search: Does It Matter?", *Journal of Labor Economics*, 36(1):1-46.

FONTAINE F. (2008) "Why Are Similar Workers Paid Differently? The Role of Social Networks", *Journal of Economic Dynamics and Control*, 32(12): 3960-3977.

FRENCH R., P. OREOPOULOS (2017) "Behavioral Barriers Transitioning to College", *Labour Economics*, 47: 48-63.

GALENIANOS M. (2013) "Learning about Match Quality and the Use of Referrals", *Review of Economic Dynamics*, 16(4): 668-690.

GALEOTTI A., L.P. MERLINO (2014) "Endogenous Job Contact Networks", *International Economic Review*, 55(4): 1201-1226.

GALOR O., J. ZEIRA (1993) "Income Distribution and Macroeconomics", *The Review of Economic Studies*, 60(1): 35-52.

GOEBEL J., M.M. GRABKA, S. LIEBIG, M. KROH, D. RICHTER, C. SCHRÖDER, J. SCHUPP (2019) "The German Socio-Economic Panel (SOEP)", *Jahrbücher für Nationalökonomie und Statistik*, 239(2): 345-360.

GLITZ A. (2017) "Coworkers Networks in the Labour Market", *Labour Economics*, 44: 218-230.

HERTZ T., T. JAYASUNDERA, P. PIRAINO, S. SELCUK, N. SMITH, A. VERESHCHAGINA (2008) "The Inheritance of Educational Inequality: Intergenerational Comparisons and Fifty-Year Trends", *The B.E. Journal of Economic Analysis and Policy*, 7(2): 1-48.

HEINECK G., R.T. RIPHAHN (2009) "Intergenerational Transmission of Educational Attainment in Germany - The Last Five Decades", *Jahrbücher für Nationalökonomie und Statistik / Journal of Economics and Statistics*, 229(1): 36-60.

HOLMLUND H., M. LINDAHL, E. PLUG (2011) "The Casual Effect of Parents' Schooling on Children's Schooling: A Comparison of Estimation Methods", *Journal of Economics Literature*, 49(3): 615-651.

HORVATH G. (2014) "On-the-Job Search and Finding a Good Job Through Social Contacts", *Journal of Theoretical Economics*, 14(1): 93-125.

IOANNIDES Y.M., A.R. SOETEVENT (2006) "Wages and Employment in A Random Social Network with Arbitrary Degree Distribution", *American Economic Review*, 96(2): 270-274.

MCPHERSON M., L. SMITH-LOVIN, J.M. COOK (2001) "Birds of a Feather: Homophily in Social Networks", *Annals Review of Sociology*, 27: 415-444.

MONTGOMERY J.D. (1991) "Social Networks and Labor-Market Outcomes: Toward an Economic Analysis", *American Economic Review*, 81: 1408-1418.

KATZ L.F., J.R. KLING, J.B. LIEBMAN (2001) "Moving To Opportunity in Boston: Early Results of a Randomized Mobility Experiment", *Quarterly Journal of Economics*, 116(2): 607-654.

KRAMARZ F., O.N. SKANS (2014) "When Strong Ties are Strong: Networks and Youth Labour Market Entry", *The Review of Economic Studies*, 81: 1164-1200.

LANG F. R., D. WEISS, A. STOCKER, B. VON ROSENBLADT (2007) "Assessing Cognitive Capacities in Computer-Assisted Survey Research: Two Ultra-Short tests of Intellectual Ability in the German Socio-Economic Panel (SOEP)", *Schmollers Jahrbuch*, 127, 183-192.

LAVECCHIA A.M., P. OREOPOULOS, R.S. BROWN (2020) "Long-Run Effects from Comprehensive Student Support: Evidence from Pathways to Education", *American Economic Review: Insights* 2(2), 209-224.

LEONE T. (2019) "The Geography of Intergenerational Mobility: Estimates of the Worldwide Variation", *Journal of Economic Development*, 44(4): 1-42.

LOHMANN H., S. WITZKE (2011) "BIOEDU (Beta Version): Biographical Data on Educational Participation and Transitions in the German Socio-Economic Panel Study (SOEP)", *Technical Report, DIW Data Documentation*

MCCRAE R.R., COSTA P.T. Jr. (1999) "A Five-Factor Theory of Personality", In: Pervin LA, John OP (Eds), *Handbook of Personality: Theory and Research*, Guilford: New York, p. 139-153.

NEUGART M., A.ZAHARIEVA (2018) "Social Networks, Promotions, and the Glass-Ceiling Effect", *Bielefeld Working Papers in Economics and Management*, No 7.

PATACCHINI E., Y. ZENOU (2011) "Neighborhood Effects and Parental Involvement in the Intergenerational Transmission of Education", *Journal of Regional Science*, 51(5): 987-1013.

PIKETTY T. (2000) "Theories of Persistent Inequality and Intergenerational Mobility", Chapter 8 of *Handbook of Income Distribution*, Volume 1: 429-476.

PLUG E., B. VAN DER KLAAUW, L. ZIEGLER (2018) "Do Parental Networks Pay Off? Linking Children's Labor-Market Outcomes to Their Parents' Friends", *The Scandinavian Journal of Economics*, 120(1): 268-295.

RIPHAHN R.T., F. SCHIEFERDECKER (2012) "The Transition to Tertiary Education and Parental Background Over Time", *Journal of Population Economics*, 25(2): 635-675.

WITTE J.C., A.L. KALLEBERG (1995) "Matching Training and Jobs: The Fit Between Vocational Education and Employment in the German Labour Market", *European Sociological Review*, 11(3): 293-317.

SACERDOTE B. (2007) "How Large are the Effects from Changes in Family Environment? A Study of Korean American Adoptees", *The Quarterly Journal of Economics*, 122(1): 119-157.

STUPNYTSKA Y., ZAHARIEVA A. (2015) "Explaining U-Shape of the Referral Hiring Pattern in a Search Model with Heterogeneous Workers", *Journal of Economic Behavior & Organization*, 119: 211-233.

STUPNYTSKA Y., ZAHARIEVA A. (2017) "Optimal Policy and The Role of Social Contacts in a Search Model with Heterogeneous Workers", *Journal of Public Economic Theory*, 19(5): 957-985.

ZAHARIEVA A. (2013) "Social Welfare and Wage Inequality in Search Equilibrium with Personal Contacts", *Labour Economics*, 23: 107-121.

ZAHARIEVA A. (2018) "On the Optimal Diversification of Social Networks in Frictional Labour Market with Occupational Mismatch", *Labour Economics*, 50: 112-127.