

Social Effects in Employer Learning: An Analysis of Siblings

Neel Rao*

University at Buffalo, SUNY

November 3, 2015

Abstract

This paper examines whether wages are based on information about personal contacts. I develop a theory of labor markets with imperfect information in which related workers have correlated abilities. I study wage setting under two alternative processes: individual learning, under which employers observe only a worker's own characteristics, and social learning, under which employers also observe those of a relative. Using sibling data from the NLSY79, I test for a form of statistical nepotism in which a sibling's performance is priced into a worker's wage. Empirically, an older sibling's test score has a larger impact on a younger sibling's log wage than a younger sibling's test score has on an older sibling's log wage. The estimates provide strong support for social effects in employer learning.

JEL Classification: C31, D83, D85, J24, J31, J71

Keywords: employer learning, human capital, sibling data, social effects, wage setting

*E-mail: neelrao@buffalo.edu. Tel: 716-645-8674. Address: 423 Fronczak Hall, Buffalo, NY 14260. I am grateful to Raj Chetty, Oliver Hart, and especially Lawrence Katz for their advice and support. I also thank Gary Chamberlain, Isaac Ehrlich, Roland Fryer, Edward Glaeser, Claudia Goldin, Susan Helper, Nishtha Langer as well as seminar participants at Case Western Reserve, FDIC, Harvard, ISB, University at Buffalo for helpful comments. Two anonymous referees and the co-editor Etienne Wasmer made suggestions that strengthened the paper.

1 Introduction

An important question in labor economics concerns how personal contacts influence job search behavior and wage setting decisions. As Granovetter’s (1974) classic survey of workers in the Boston area illustrates, nearly half of all jobs are obtained through a social contact. The extensive use of friends, relatives, and acquaintances in job search enables personal contacts to play a role in shaping employers’ beliefs about a worker’s skills. As Rees (1966) notes when studying workers in a Chicago neighborhood, “Present employees tend to refer people like themselves, and they may feel that their own reputation is affected by the quality of the referrals.”

This paper develops an empirical test for whether wages incorporate information on personal contacts. Combining a sibling model similar to Griliches (1979) with an employer learning model related to Altonji and Pierret (2001), I construct a framework in which workers are organized into disjoint social groups composed of agents with correlated abilities and differing ages. I examine wage determination under two competing assumptions about the market’s formation of beliefs: individual learning and social learning. Under individual learning, a worker’s wage equals the conditional expectation of her productivity given only her own schooling and performance. Under social learning, a worker’s wage equals the conditional expectation of her productivity given the schooling and performance of all the members of her social group, including herself.

Using sibling data from the NLSY79, I apply this framework to test for a form of statistical nepotism in which a worker’s wage depends on both one’s own and a sibling’s characteristics.¹ The basic logic is as follows. If one sibling is older than another sibling, then employers should have more precise information about the older sibling. Hence, when market participants form Bayesian beliefs about the abilities of the two siblings, the older sibling’s average performance would have a greater impact on employers’ mean beliefs about the younger sibling’s ability than vice versa, and the component of the younger sibling’s wage attributable to the older sibling’s ability would be larger than the component of the older sibling’s wage attributable to the younger sibling’s ability.

Empirically, given data on the test scores and schooling of siblings, this weighting can be detected by regressing an individual’s log wage on her own and a sibling’s test scores and schooling. If employer learning is nepotistic, then the ratio of the coefficient on a sibling’s test score to that on one’s own test score should typically be higher in a younger sibling’s than in an older sibling’s log wage. However, if employer learning is individual, then the ratio of the coefficient on a sibling’s test score to that on one’s own test score should be the same for both a younger and an older sibling. The empirical results are consistent with statistical nepotism. In order to eliminate other explanations for the results, I document several pieces of evidence related to job search patterns, human capital measures, and geographic or economic proximity.

¹This article uses the term nepotism to refer to unequal treatment on the basis of family relationships. This usage differs from the nomenclature in Becker (1971), who applies the term to describe favoritism towards a particular group. By statistical nepotism, I mean unequal treatment due to information from a relative, not because of the preferences or influence of a relative.

The empirical strategy integrates elements from four largely distinct literatures in labor economics. First, this paper is part of a sizeable literature on the identification of social effects.² The most closely related paper is Case and Katz (1991), which attempts to detect neighborhood influences by regressing an individual’s outcome variable on the background variables of her peers.

Second, this paper contributes to a growing literature on employer learning and statistical discrimination. In order to examine social interactions in the employer learning process, I extend the basic methodology developed by Farber and Gibbons (1996) and Altonji and Pierret (2001).³ I thereby devise a test for statistical nepotism, in which employers infer an individual’s productivity based partly on information about her relatives.

Third, this paper is relevant to a theoretical literature on social networks in labor markets.⁴ The framework in the current paper is most similar to the model in Montgomery (1991). In that model, workers are arranged into social groups containing either one or two members, and social groups of size two consist of an older and a younger worker with correlated abilities. Employers use the observed performance of the more senior worker in each pair to infer the ability of her more junior counterpart.

Fourth, this paper contributes to a small empirical literature that attempts to test for nepotism in labor markets. A relevant paper is Lam and Schoeni (1993), whose empirical strategy involves comparing the coefficients on a father’s and a father-in-law’s schooling in wage equations.⁵ In addition, Wang (2013) identifies the effect of a father-in-law on male earnings, and Gevrek and Gevrek (2010) study how nepotism influences college performance.

The remainder of this paper is structured as follows. Section 2 presents the employer learning models. Section 3 describes the empirical specification. Section 4 discusses the construction of the estimation sample. Section 5 presents the empirical results. Section 6 proposes some implications for antidiscrimination policy. Section 7 concludes.

2 Sibling Models with Employer Learning

This section analyzes how employer learning affects the relationship between siblings’ test scores and log wages. The framework embeds a sibling model based on Griliches (1977, 1979) into an employer learning model related to Farber and Gibbons (1996), Altonji and Pierret (2001), and Lange (2007).

²See Ioannides and Loury (2004) for a review of existing research on social effects in labor markets. Recent papers in this area include Bayer et al. (2008), Beaman (2012), Beaman and Magruder (2012), Burns et al. (2010), Cingano and Rosolia (2012), Combes et al. (2008), Dustmann et al. (2011), Hellerstein et al. (2011), Kramarz and Skans (2013), Kugler (2003), and Nakajima et al. (2010).

³Pinkston (2003) analyzes gender differences in the process of employer learning. Autor and Scarborough (2008) investigate an example of statistical discrimination based on race. Simon and Warner (1992) as well as Pinkston (2012) assess the informational content of employee referrals.

⁴Models of job search through social networks have been developed by Bramoullé and Saint-Paul (2010), Calvò-Armengol and Jackson (2004), Mortensen and Vishwanath (1994), and Zaharieva (2013).

⁵Similarly, Hellerstein and Morrill (2011) examine trends in the transmission of human capital from fathers to daughters by analyzing changes in the likelihood that a woman enters her father’s as compared to her father-in-law’s occupation.

2.1 Labor Market Characteristics of Siblings

This section presents a statistical model of siblings' labor market attributes. The treatment focuses on the case in which there are two siblings, 1 and 2. As in much of the literature on employer learning, the log labor productivity $l(s_i, a_i, t_i)$ of person $i \in \{1, 2\}$ is assumed to be decomposable into two components:

$$l(s_i, a_i, t_i) = g(s_i, a_i) + h(t_i), \quad (1)$$

where $g(s_i, a_i)$ is a time-invariant component of productivity, and $h(t_i)$ represents additional human capital accumulated with age t_i . Letting $\beta > 0$, the function $g(s_i, a_i)$ is linear in schooling s_i and ability a_i :

$$g(s_i, a_i) = \beta s_i + a_i, \quad (2)$$

where the coefficient on a_i is without loss of generality normalized to one.

The abilities a_1, a_2 of the two siblings are joint normally distributed with respective means μ_{a1} and μ_{a2} , identical variance $\sigma_a^2 > 0$, and correlation $\rho_a \in (0, 1)$. Letting $\gamma > 0$, schooling is related to ability through:

$$s_i = \gamma a_i + \epsilon_i, \quad (3)$$

where ϵ_i , which represents factors other than labor market ability that influence education decisions, is assumed to be independent of a_1, a_2 . The error terms ϵ_1, ϵ_2 are joint normally distributed with respective means $\mu_{\epsilon 1}$ and $\mu_{\epsilon 2}$, identical variance $\sigma_\epsilon^2 > 0$, and correlation $\rho_\epsilon \in (0, 1)$.

The information available to employers about ability a_i is symmetric but imperfect.⁶ In particular, employers observe the schooling s_i of each person as well as a sequence $r_i = \{r_{iu}\}_{u=1}^{t_i}$ of noisy productivity signals given by:

$$r_{iu} = g(s_i, a_i) + \eta_{iu}, \quad (4)$$

where each measurement error η_{iu} is a normal random variable with mean zero and variance $\sigma_\eta^2 > 0$.⁷ The η_{iu} are assumed to be independent of each other and of all the other variables in the model.

The econometrician is assumed to observe a test score z_i in addition to the education level s_i . Letting $\theta_s > 0$ and $\theta_a > 0$, the ability measure z_i takes the form:

$$z_i = \theta_s s_i + \theta_a a_i + \omega_i, \quad (5)$$

where ω_i , which represents factors unrelated to labor productivity that affect the test score, is independent of both a_1, a_2 and s_1, s_2 . The error terms ω_1, ω_2 are joint normally distributed with respective means $\mu_{\omega 1}$ and $\mu_{\omega 2}$, identical variance $\sigma_\omega^2 > 0$, and correlation $\rho_\omega \in (0, 1)$. In addition, the test score z_i is assumed to be unobservable to employers as in Altonji and Pierret (2001); so that, employers cannot

⁶Symmetric information means that all employers in the labor market are equally knowledgeable about the variables in the model. That is, prospective employers observe the same information about a worker as a current employer.

⁷Note that r_i can be interpreted as the performance history of a worker.

use z_i as an additional signal of productivity when forming beliefs about a_i .

2.2 Individual Learning

This section analyzes the case where employer learning is individualistic; so that, the wage w_i of person $i \in \{1, 2\}$ is based only on her own education s_i and her own performance r_i . For concreteness, employers are assumed to set the wage w_i equal to the conditional expectation of labor productivity given s_i and r_i . The analysis here proceeds in two steps. I first express $\log(w_i)$ as a function of s_i and r_i . I then examine the conditional expectation of $\log(w_i)$ given s_1, s_2 and z_1, z_2 .

In order to calculate $\log(w_i)$, I derive beliefs given s_i and r_i .⁸ Conditional on schooling s_i , employers' beliefs about $g(s_i, a_i)$ are normally distributed with mean $\mu_{mi}(s_i)$ and variance σ_m^2 where:

$$\mu_{mi}(s_i) = \mathbb{E}[g(s_i, a_i)|s_i] = \beta s_i + \mathbb{E}(a_i|s_i) \quad \text{and} \quad \sigma_m^2 = \mathbb{V}[g(s_i, a_i)|s_i] = \mathbb{V}(a_i|s_i). \quad (6)$$

From the results in DeGroot (1970), it follows that employers' beliefs about $g(s_i, a_i)$ given both s_i and r_i are normally distributed with mean $\mu_{gi}(s_i, r_i)$ and variance σ_{gi}^2 where:

$$\mu_{gi}(s_i, r_i) = (1 - \chi_i)\mu_{mi}(s_i) + \chi_i\bar{r}_i, \quad \sigma_{gi}^2 = (\sigma_m^{-2} + t_i\sigma_\eta^{-2})^{-1}, \quad \chi_i = t_i\sigma_\eta^{-2}\sigma_{gi}^2, \quad (7)$$

and \bar{r}_i is the sample mean of the sequence $r_i = \{r_{iu}\}_{u=1}^{t_i}$. If worker 1 is older than worker 2, then χ_1 will be larger than χ_2 , indicating that beliefs about worker 1 are based less on schooling and more on performance relative to beliefs about worker 2.

Given the normality of employers' beliefs about $g(s_i, a_i)$, the conditional expectation of labor productivity given s_i and r_i can be expressed as:

$$\mathbb{E}\{\exp[l(s_i, a_i, t_i)]|s_i, r_i\} = \exp[\mu_{gi}(s_i, r_i) + \frac{1}{2}\sigma_{gi}^2 + h(t_i)]; \quad (8)$$

so that, the log wage of each person is simply:

$$\log(w_i) = \mu_{gi}(s_i, r_i) + \frac{1}{2}\sigma_{gi}^2 + h(t_i), \quad (9)$$

where $\mu_{gi}(s_i, r_i)$ and σ_{gi}^2 are given by equation (7).

I now calculate the conditional expectation of the log wage given the information available to the econometrician. Using equations (4), (7), and (9), one obtains:

$$\mathbb{E}[\log(w_i)|s_1, s_2, z_1, z_2] = \chi_i\mathbb{E}(a_i|s_1, s_2, z_1, z_2) + f_i(s_i, t_i), \quad (10)$$

⁸The notation $\mathbb{E}(\cdot)$ and $\mathbb{V}(\cdot)$ is used for the unconditional expectation and variance. The conditional expectation and conditional variance are respectively denoted by $\mathbb{E}(\cdot|\cdot)$ and $\mathbb{V}(\cdot|\cdot)$.

where the function $f_i(s_i, t_i)$ is given by:

$$f_i(s_i, t_i) = (1 - \chi_i)\mu_{mi}(s_i) + \chi_i\beta s_i + \frac{1}{2}\sigma_{gi}^2 + h(t_i). \quad (11)$$

The following is an invariance result concerning the ratio of the coefficients on the test scores in a regression of $\log(w_1)$, $\log(w_2)$ on s_1 , s_2 and z_1 , z_2 . This property is an immediate consequence of equation (10).

In particular, this implication of the model arises from the symmetric treatment of siblings. The specified relationships among ability, schooling, and test scores are similar for younger and older siblings. In equation (10), the two siblings have the same coefficients as each other on one's own and a sibling's test scores and schooling in the conditional expectation of one's ability. These equalities follow from the formulae for schooling and test scores in equations (3) and (5). The function relating the test score to schooling and ability, as well as that relating schooling to ability, is the same for both siblings.

Proposition 1 *Suppose that learning is individual. Let α_{ij} denote the regression coefficient on person j 's test score in the conditional expectation of person i 's log wage given s_1 , s_2 and z_1 , z_2 . Then $\alpha_{12}\alpha_{22} = \alpha_{21}\alpha_{11}$.*

To understand this result, suppose that sibling 1 is older than sibling 2; so that, sibling 1's wage is based less on education and more on performance than sibling 2's wage. Because each sibling's wage has a different composition, it might be difficult to compare the results of wage regressions across siblings. The importance of proposition 1 is that it enables such comparisons to be made. Even though α_{11} is larger in magnitude than α_{22} by the proportion χ_1/χ_2 , it follows from proposition 1 that α_{12} is also larger in magnitude than α_{21} by this proportion. In other words, the impact of a sibling's test score on one's log wage grows with age at the same rate as the impact of one's own test score. Because employers do not use information on a person's siblings under individual learning, this result is valid regardless of the number of siblings in a family. Section 2.3 examines how deviations from this rule can arise when employers use information on one sibling when determining the wage of another sibling.

The coefficient restriction above follows from the symmetric modeling of siblings' human capital. Given that older and younger siblings have similar underlying skills, asymmetries in log wage regressions may be attributed to social effects in employer learning. Nonetheless, there are several mechanisms that might lead to differences between siblings in the parameters regulating skill formation. The potential effects of these mechanisms on the specification of the model are discussed below. In the empirical analysis, a number of tests are conducted in order to demonstrate that such alternative explanations are not critical for generating the patterns observed in the data.

One possibility is that the process of human capital development in early life varies with birth order. If so, there may be differences between siblings in how labor productivity, schooling, and test scores are modeled in equations (2), (3), and (5). Specifically, the parameters θ_a and θ_s representing the effects of ability and schooling on test scores may be a function of birth order as well as variables like birth

cohort, sibship size, and birth spacing. The same applies to the coefficient β , which embodies the effect of schooling on earnings, and the parameter γ , which captures the influence of ability on schooling. Furthermore, as described below, the equations specifying a person's skills may depend on a sibling's characteristics as well as one's own. See Cunha and Heckman (2007) for a model of skill production in childhood and parental investments in offspring.

Another sort of interaction involves transfers of skills between siblings, either as children in the parental home or as adults in the labor market. For example, Zajonc (1976) notes that older children might serve as teachers for younger children. Another type of peer influence is a role model effect, whereby the actions of one sibling provide signals of appropriate behavior for another sibling to follow. For example, Butcher and Case (1994) mention that individuals might compare their own achievements with those of their siblings when making educational decisions. In such situations, a person's schooling may depend not only on one's own ability but also on a sibling's ability and schooling. Similarly, both one's own and a sibling's ability and schooling may affect one's labor productivity and test score. That is, a person's skills as formulated in equations (2), (3), and (5) could be functions of both one's own and a sibling's attributes. In addition, these functions might differ between older and younger siblings.

Under individual learning, the parameters α_{11} and α_{22} are predicted to be strictly positive.⁹ Accordingly, the equality in proposition 1 can be expressed as $\alpha_{21}/\alpha_{22} = \alpha_{12}/\alpha_{11}$. That is, the ratio of the coefficient on a sibling's test score to the coefficient on one's own test score does not differ between siblings in the regression of one's log wage on one's own and a sibling's schooling and test scores. For expositional purposes, the ratios α_{21}/α_{22} and α_{12}/α_{11} will each be called a scaled sibling effect, which is abbreviated as SSE. In addition, the restriction $\alpha_{21}/\alpha_{22} = \alpha_{12}/\alpha_{11}$ will be termed equal scaled sibling effects, whose acronym is ESSE.

2.3 Social Learning

This section examines the case in which employer learning has an element of statistical nepotism. In particular, the wage w_i of each sibling incorporates information on the education s_1, s_2 and performance r_1, r_2 of both siblings; so that, employers set the wage of sibling $i \in \{1, 2\}$ equal to the conditional expectation of her own labor productivity given s_1, s_2 and r_1, r_2 .¹⁰ I first derive the log wage as a function of the information available to employers and then compute the conditional expectation of the log wage given the variables observable to the econometrician.

I begin by calculating beliefs given s_1, s_2 and r_1, r_2 . For a given sibling with index i , let e be the

⁹This property follows from the results in the appendix, where the regression coefficient on one's own test score is shown to be positive in the conditional expectation of one's ability given one's own and a sibling's schooling and test scores.

¹⁰A potential issue is that a sibling's characteristics may not be observable to a person's employer unless both individuals work at the same firm. To address this point, the online appendix presents a simple model of employee referrals in which an older sibling's attributes can affect a younger sibling's log wage even if the two siblings work for different employers. In the model, a person refers a sibling to her employer, who then makes a wage offer to the sibling. The wage offer can act as a signal to other firms of the sibling's ability even if the sibling does not accept the offer.

index of the other sibling; so that, $e = 2$ if $i = 1$, and vice versa. Conditional on the schooling s_i, s_e of both siblings and the performance r_e of sibling e , employers' beliefs about the time-invariant component $g(s_i, a_i)$ of sibling i 's log productivity are normally distributed with mean $\mu_{ni}(s_i, s_e, r_e)$ and variance σ_{ni}^2 where:

$$\mu_{ni}(s_i, s_e, r_e) = \beta s_i + \mathbb{E}(a_i | s_i, s_e, r_e) \text{ and } \sigma_{ni}^2 = \mathbb{V}(a_i | s_i, s_e, r_e). \quad (12)$$

Note that the conditional variances $\sigma_{ni}^2, \sigma_{ne}^2$ satisfy $\sigma_{ni}^2 \geq \sigma_{ne}^2$ if $t_i \geq t_e$. In other words, if sibling i is at least as old as sibling e , then beliefs about sibling e conditional only on s_e, s_i , and r_i are at least as precise as beliefs about sibling i conditional only on s_i, s_e , and r_e .

Employers' beliefs about $g(s_i, a_i)$ given both s_i, s_e and r_i, r_e are normally distributed with mean $\mu_{qi}(s_i, s_e, r_i, r_e)$ and variance σ_{qi}^2 where:

$$\mu_{qi}(s_i, s_e, r_i, r_e) = (1 - \xi_i)\mu_{ni}(s_i, s_e, r_e) + \xi_i \bar{r}_i, \quad \sigma_{qi}^2 = (\sigma_{ni}^{-2} + t_i \sigma_\eta^{-2})^{-1}, \quad \xi_i = t_i \sigma_\eta^{-2} \sigma_{qi}^2, \quad (13)$$

and \bar{r}_i is the sample mean of $\{r_{iu}\}_{u=1}^{t_i}$. In equation (13), if t_i is greater than t_e , then ξ_i is greater than ξ_e . To paraphrase, if sibling i is older than sibling e , then beliefs about sibling i are based more on her own performance and less on other measures of her ability compared to beliefs about sibling e .

It follows that the conditional expectation of sibling i 's labor productivity given s_i, s_e and r_i, r_e is:

$$\mathbb{E}\{\exp[l(s_i, a_i, t_i)] | s_i, s_e, r_i, r_e\} = \exp[\mu_{qi}(s_i, s_e, r_i, r_e) + \frac{1}{2}\sigma_{qi}^2 + h(t_i)], \quad (14)$$

resulting in the proceeding expression for sibling i 's log wage:

$$\log(w_i) = \mu_{qi}(s_i, s_e, r_i, r_e) + \frac{1}{2}\sigma_{qi}^2 + h(t_i), \quad (15)$$

where $\mu_{qi}(s_i, s_e, r_i, r_e)$ and σ_{qi}^2 are given by equation (13).

I now derive the conditional expectation of $\log(w_i)$ given s_i, s_e and z_i, z_e . Combining equations (4), (12), (13), and (15), one obtains:

$$\mathbb{E}[\log(w_i) | s_i, s_e, z_i, z_e] = (1 - \xi_i)\mathbb{E}[\mathbb{E}(a_i | s_i, s_e, r_e) | s_i, s_e, z_i, z_e] + \xi_i \mathbb{E}(a_i | s_i, s_e, z_i, z_e) + b_i(s_i, t_i), \quad (16)$$

where the function $b_i(s_i, t_i)$ is given by:

$$b_i(s_i, t_i) = \beta s_i + \frac{1}{2}\sigma_{qi}^2 + h(t_i). \quad (17)$$

Because the r_{ue} have identical covariances with each other and with s_i, s_e , and a_i , the conditional expectation of a_i given s_i, s_e , and r_e has the form:

$$\mathbb{E}(a_i | s_i, s_e, r_e) = \mathbb{E}(a_i | s_i, s_e, \bar{r}_e) = \zeta_{ci} + \zeta_{oi}s_i + \zeta_{fi}s_e + \zeta_{ri}\bar{r}_e; \quad (18)$$

so that, (s_i, s_e, \bar{r}_e) is a sufficient statistic for (s_i, s_e, r_e) with respect to a_i . Thus, the iterated expectation in equation (16) can be expressed as:

$$\mathbb{E}[\mathbb{E}(a_i | s_i, s_e, r_e) | s_i, s_e, z_i, z_e] = \zeta_{ri} \mathbb{E}(a_e | s_i, s_e, z_i, z_e) + d_i(s_i, s_e), \quad (19)$$

where $d_i(s_i, s_e)$ is defined as:

$$d_i(s_i, s_e) = \zeta_{ci} + \zeta_{oi} s_i + (\zeta_{fi} + \zeta_{ri} \beta) s_e. \quad (20)$$

Combining equations (16) and (19), I obtain the final expression for the conditional expectation of the log wage:

$$\mathbb{E}[\log(w_i) | s_i, s_e, z_i, z_e] = (1 - \xi_i) \zeta_{ri} \mathbb{E}(a_e | s_i, s_e, z_i, z_e) + \xi_i \mathbb{E}(a_i | s_i, s_e, z_i, z_e) + p_i(s_i, s_e, t_i), \quad (21)$$

where $p_i(s_i, s_e, t_i)$ is given by:

$$p_i(s_i, s_e, t_i) = b_i(s_i, t_i) + (1 - \xi_i) d_i(s_i, s_e). \quad (22)$$

Equation (21) demonstrates that the log wage can be decomposed into two separate components, one of which contains a person's own ability, and the other of which reflects her sibling's ability.

It is now possible to prove the following counterpart to proposition 1 for the current model in which employer learning has a social component.¹¹ The first part of the proposition is an immediate consequence of the symmetric treatment of the two siblings. In the second part, an analogous statement holds if $t_2 > t_1$ instead of $t_1 > t_2$.

Proposition 2 *Suppose that learning is social. Let ν_{ij} denote the regression coefficient on person j 's test score in the conditional expectation of person i 's log wage given s_1, s_2 and z_1, z_2 .*

1. *If $t_1 = t_2$, then $\nu_{12}\nu_{22} = \nu_{21}\nu_{11}$.*
2. *If $t_1 > t_2$, then $\nu_{12}\nu_{22} < \nu_{21}\nu_{11}$.*

On the one hand, the first part of proposition 2 is a variant of the results in Manski (1993) concerning the difficulties of distinguishing between social and nonsocial effects. If the two siblings have the same age, then both the individual and the social learning model predict that the ratio of the coefficients on test scores should be the same for the two siblings. On the other hand, if there are asymmetries in the ages of the siblings, then the two models generate different predictions regarding the relative values of this ratio, making it possible to detect social effects in employer learning.

Intuitively, if sibling 1 is older than sibling 2, then employers have more precise information about sibling 1 than about sibling 2, because sibling 1's performance has been observed for longer than sibling

¹¹The proof of the proposition is given in the appendix.

2's. Thus, when the market forms Bayesian beliefs about the abilities of the two siblings, greater weight is placed on information about sibling 1 than on information about sibling 2. Given that the labor market is competitive, this relative weighting is reflected in the wages of the two siblings; so that, the component of sibling 2's wage attributable to sibling 1's ability is larger than the component of sibling 1's wage attributable to sibling 2's ability.

This phenomenon manifests itself in the data available to the econometrician as follows. If sibling 1 is older than sibling 2, then sibling 2's log wage is more strongly influenced by sibling 1's ability than vice versa. Moreover, a person's ability is more closely associated with one's own than with a sibling's test score. Consequently, the ratio of the coefficient on a sibling's test score to the coefficient on one's own test score is typically higher in sibling 2's log wage than in sibling 1's log wage.

The validity of this test relies on a symmetry between the skills of siblings. In order to interpret an asymmetric wage structure as being due to social effects in employer learning, other types of peer influence between siblings must be limited. However, the model does not formally account for mechanisms such as transfers of human capital, interactions in skill development, and role model effects. As previously explained, these factors could cause dissimilarities between siblings in the specifications of labor productivity, schooling, and test scores. Therefore, several exercises are performed with the data so as to isolate social learning as the most plausible explanation for the findings. In particular, the relationship among underlying measures of human capital is analyzed, and estimates are computed for siblings working in the same field or living in different regions. The role of job tenure is also examined.

Finally, suppose that $t_1 > t_2$, meaning that sibling 1 is older than sibling 2. If the parameters ν_{11} and ν_{22} are strictly positive, then the inequality in proposition 2 can be formulated as $\nu_{21}/\nu_{22} > \nu_{12}/\nu_{11}$. That is, the ratio of the coefficient on an older sibling's test score to that on one's own test score in the conditional expectation of a younger sibling's log wage is greater than the ratio of the coefficient on a younger sibling's test score to that on one's own test score in the conditional expectation of an older sibling's log wage. In other words, the scaled sibling effect on a younger sibling's log wage is larger than the scaled sibling effect on an older sibling's log wage. This is a deviation from the principle of equal scaled sibling effects. Mnemonically speaking, ESSE is violated because the SSE on a younger sibling is bigger than the SSE on an older sibling.

3 Empirical Implementation

The following is the basic strategy for estimating and testing the individual and social learning models.¹² First, the log wages of all older siblings across all years are regressed both on their own and their younger sibling's test scores, schooling, and background attributes and on a bivariate polynomial in their own and their younger sibling's ages. Second, the log wages of all younger siblings across all years are regressed both on their own and their older sibling's test scores, schooling, and background attributes and on a bivariate polynomial in their own and their older sibling's ages. Third, to evaluate the nonlinear

¹²See the online appendix for additional technical details regarding this procedure.

restriction implied by the individual learning model, I calculate the Wald statistic for the null hypothesis that the coefficient on a younger sibling’s test score in an older sibling’s log wage times the coefficient on one’s own test score in a younger sibling’s log wage minus the coefficient on an older sibling’s test score in a younger sibling’s log wage times the coefficient on one’s own test score in an older sibling’s log wage is equal to zero.¹³ When computing standard errors and test statistics, the Huber-White estimator of the variance-covariance matrix is used to allow for arbitrary forms of correlation among the error terms of observations on the same family.

4 Dataset Construction and Description

The dataset is constructed from the 1979-2008 waves of the National Longitudinal Survey of Youth 1979 (NLSY79), which contains panel data on 12,686 men and women aged 14-22 in 1979. Respondents were interviewed annually from 1979 to 1994 and biennially thereafter. The NLSY79 is especially well suited to the purpose of this paper. Because the Armed Services Vocational Aptitude Battery (ASVAB) was administered to participants in the NLSY79, a growing literature on employer learning uses the resulting Armed Forces Qualification Test (AFQT) score as an ability measure that is not directly observable to employers. In addition, a large number of sibling studies analyze data from the NLSY79, which includes 5,863 respondents who have one or more interviewed siblings.

In order to implement the empirical strategy, I assemble a dataset in which each observation represents a particular sibling pair in a given survey year.¹⁴ This dataset will serve as the main estimation sample for the paper; therefore, the current section describes in detail how this dataset is constructed.¹⁵ The data are derived from the 6,111 respondents in the cross-sectional sample and the 5,295 respondents in the supplemental sample of the NLSY79. I identify every survey year in which a respondent has a non-missing wage observation on a full-time job, where full-time is defined as 35 or more hours per week. Each wage is then deflated using the CPI to a base period of 1982-1984, and any real hourly wage less than \$1 or greater than \$100 is omitted from the analysis.

I next compile information on each respondent’s education and AFQT score. The AFQT scores are standardized among all respondents in the NLSY79 with the same year of birth. I exclude all observations that occur prior to the first survey year in which a respondent has left school for the first time. At this point, each observation corresponds to a specific person in a given survey year.

To generate a new dataset consisting of sibling pairs instead of individuals, I apply the following

¹³By the delta method, this test statistic is, in general, asymptotically distributed as chi-squared with one degree of freedom.

¹⁴If there are two siblings p and q , then a sibling pair in which sibling p appears first and sibling q second is regarded as distinct from a sibling pair in which sibling q appears first and sibling p second. For example, if a family consists of three siblings, then six different pairs of siblings can be formed.

¹⁵The main estimation sample is restricted to observations on sibling pairs in which both members have worked since the last interview. However, selection into employment may not be entirely exogenous. To address this issue, the dataset was expanded to include non-working individuals, and the joint work-wage outcomes of respondents were examined. I continue to find evidence of sibling effects on labor market earnings after performing this extension. These results are available in the online appendix.

procedure. For each person in the sample in a given survey year, I search over all the other individuals in the sample in that year for an observation on the person’s sibling. Whenever a sibling is found, an observation containing information on the person and her sibling is added to the new dataset in that survey year. Because the empirical strategy is based primarily on age differences, any pairs of siblings having the same year and month of birth are eliminated. The resulting dataset contains 54,474 observations on 7,074 sibling pairs, covering 4,726 individuals from 1,993 families. Table 1 presents descriptive statistics.

5 Empirical Results

This section presents empirical evidence of social interactions among siblings in the labor market.

5.1 Job Search Estimates

Job search is a channel through which employers can acquire information about a person’s siblings.¹⁶ The first four columns in Table 2 estimate linear probability models relating birth order and sibship size to the likelihood of obtaining one’s most recent job with the help of a sibling. When both sibship size and birth order are jointly included in the regression, the coefficient on sibship size is not significantly different from zero, whereas the coefficient on birth order is statistically significant. This finding is essentially unchanged after controlling for a variety of additional demographic variables. The second four columns present estimates of the impact of birth order and sibship size on the likelihood that an individual obtained her most recent job with the help of a sibling who was working for the employer that offered her the job. The pattern of results in the second four columns is largely similar.¹⁷

5.2 Sibling AFQT Impacts

This section presents the main estimates of sibling effects. Table 3 displays the impacts of one’s own and a sibling’s AFQT scores on a person’s log wage. I differentiate between the two employer learning models based on the predictions in propositions 1 and 2. The SSE on a younger sibling is significantly greater than the SSE on an older sibling. This finding contradicts the individual learning model in proposition 1 but is consistent with the social learning model in proposition 2. An older sibling’s ability seems to have a larger influence on a younger sibling’s log wage than vice versa.

¹⁶Detailed summary statistics for the role of different relatives in job search are available in the online appendix. In the sample here, 4.52 percent of individuals found their most recent employer with the help of a sibling, and 3.68 percent of individuals report that this sibling also worked for their most recent employer. When tabulated by birth order, the former percentage ranges from 1.34 for first-born children to 10.35 for seventh- or later-born children, and the latter percentage ranges from 1.05 for first-born children to 8.17 for seventh- or later-born children.

¹⁷A potential issue is that the percentages of individuals obtaining a job through a sibling or also working at the same firm as a sibling might be too small to produce substantial estimates of sibling effects. To address this point, the online appendix constructs a simple model of employee referrals that can generate sibling effects even if siblings work at different firms in equilibrium. In the model, a person refers a sibling to her employer, who then makes a wage offer to the sibling. The wage offer can act as a signal to other firms of the sibling’s ability even if the sibling does not accept the offer.

Table 4 compares siblings who were residing in different geographic regions towards the beginning of their careers.¹⁸ If the empirical patterns are due to social effects in employer learning, then the observed asymmetries should largely be absent when studying younger and older siblings who have worked in different regional labor markets since the first time they were living on their own.¹⁹ The estimates fail to reject the null hypothesis of ESSE. There is essentially no evidence of differences between the impacts of the AFQT scores of younger and older siblings who were residing in different regions during the early stages of their careers.²⁰

Table 5 shows how the estimates vary depending on whether siblings work in the same or different occupations or industries.²¹ If labor market interactions are generating the documented sibling effects, then the findings should strengthen among siblings working in the same occupation or industry, because such individuals would be more likely to have come into contact with each other's employer. The upper panel displays results for siblings working in the same occupation, industry, and occupation or industry. The results strongly reject the null hypothesis of ESSE. The lower panel displays results for siblings who have always been working in different occupations, industries, and occupations and industries.²² The null hypothesis cannot be rejected in this case. There is virtually no evidence of an asymmetry between the impacts of an older and a younger sibling's AFQT scores.²³

Another question concerns changes in the wage structure with tenure at a firm. On the one hand, employers may know little about the productivity of a new hire, and so they might rely on information about a sibling when predicting ability and formulating wages. This effect may be especially relevant for siblings working in the same occupation or industry because firms may be more familiar with the sibling of a worker in such a situation. On the other hand, employers should be better informed about the capabilities of a senior employee, in which case information about one's own performance should be a stronger determinant of wages than information on a sibling. If siblings are working in the same occupation or industry, then a sibling's test score should have a large impact on the wages of new hires

¹⁸The regions above are the four Census geographic regions of the United States: Northeast, Midwest, South, and West.

¹⁹Another exercise involves estimating the relationship of the log wage to the AFQT score of a sibling with little work experience. The outcome of this falsification test is reported in the online appendix. The SSE on a younger sibling does not differ significantly from the SSE on an older sibling.

²⁰The analysis can be extended by studying how the SSE changes when two siblings initially living in the same region become geographically separated. The online appendix documents the results of this extension. The SSE decreases if either of two siblings currently living in the same region moves to a different region. This effect is especially pronounced for individuals that leave for a new job.

²¹The 2000 Census 3-digit codes for the occupation and industry of each job are used. In order to protect the privacy of survey respondents, the NLSY79 does not contain precise information on employer identity, and so researchers cannot distinguish exactly those siblings working for the same employer.

²²In order to perform this exercise, siblings should be working in different fields throughout their careers. Otherwise, if a pair of siblings initially works in the same occupation or industry, then some information on one sibling's performance might get incorporated into the other sibling's wage, and such information might be transmitted to future employers in a different field if the wage is publicly observable.

²³Estimates were also computed for siblings working in the same occupation or industry and living in the same region as well as for siblings working in different occupations and industries and living in different regions. The null hypothesis is firmly rejected for the former group but safely retained for the latter group. These results are available in the online appendix.

but only a weak relationship with the earnings of senior employees. If siblings are employed in different occupations and industries, then a sibling’s test score should have a relatively small influence on labor income irrespective of seniority.

This test is implemented in Table 6, which exhibits the impacts of one’s own and a sibling’s AFQT scores on the log wages of new hires and senior employees. A new hire is defined as an individual who has worked at her current employer for less than three months, and a worker with over three years of tenure on the job is labeled as a senior employee.²⁴ For siblings in different occupations and industries, the coefficient on a sibling’s AFQT score is not significantly different from zero, whereas the coefficient on one’s own AFQT score is highly significant. The estimates are consistent with the null hypothesis that the SSE is the same for new hires and senior employees. For siblings in the same occupation or industry, the estimated coefficient on a sibling’s AFQT score is greater than the estimated coefficient on one’s own AFQT score among new hires, whereas the opposite holds among senior employees. The null hypothesis is rejected in this sample. Hence, sibling effects seem to fade out with job tenure.

In most specifications, the coefficient on a younger sibling’s AFQT score in an older sibling’s log wage is negative, although not statistically significant. There are also some cases in which an older sibling’s AFQT score has an insignificantly negative coefficient in a younger sibling’s log wage. This finding should not be interpreted as a negative causal effect of a sibling’s performance on a person’s earnings. Instead, the sibling model in section 2.1 provides an explanation. A negative coefficient on a sibling’s test score is attributable to a negative partial correlation between one’s own ability and a sibling’s test score given one’s own schooling, a sibling’s schooling, and one’s own test score.²⁵ As described in the appendix, this situation arises from a strong correlation between siblings in the component of test scores that is unrelated to labor productivity. It occurs when the parameter ρ_ω is relatively large, in which case the error terms ω_1 and ω_2 in equation (5) are highly correlated between siblings.

Although social effects in employer learning are a plausible explanation for the observed asymmetries, other mechanisms could give rise to similar patterns. One possibility is that the results do reflect labor market interactions between siblings but that an informational mechanism such as employer learning is not involved. In particular, Becker (1971) studies taste-based discrimination, and Lam and Schoeni (1993) analyze skills-based nepotism. Specifically, employers may exhibit favoritism towards the younger sibling of a skilled older employee, and an older individual may use her ability and power within a firm to assist a younger sibling. However, such behavior does not seem to provide a satisfactory explanation. If nepotistic returns to a sibling’s human capital were driving the results, then an older sibling’s schooling should have a higher coefficient than a younger sibling’s schooling in a log wage equation. However, I

²⁴Small changes in the thresholds for being categorized as a new hire or a senior employee do not affect the basic pattern of results.

²⁵See the appendix for more details. A negative partial effect of a sibling’s test score on a person’s log wage implies that the parameter π_f is negative in equation (26). That is, the coefficient on a sibling’s test score is negative in the hypothetical regression of a person’s ability on one’s own and a sibling’s test scores and schooling. This result is compatible with the theoretical prediction from lemma 2 that π_f is smaller than π_o in absolute value. That is, the magnitude of the coefficient on a sibling’s test score is less than the magnitude of the coefficient on one’s own test score in the aforementioned regression.

obtain the opposite result in Table 3. Moreover, the fadeout with tenure of the coefficient on a sibling's AFQT score in Table 6 suggests a learning process.

In addition, several mechanisms other than labor market interactions could produce the observed asymmetries. First, there might be transfers of skills between siblings, especially while siblings live together in the parental home. Second, even if siblings do not directly transfer skills to each other, the process of skill formation in childhood might give rise to asymmetries in the relationship of one's skills to the abilities of a younger and an older sibling. Third, an older sibling might serve as a role model for a younger sibling.

Nonetheless, these three factors are unrelated to wage setting in particular and would likely affect other skills measures. Specifically, if interactions among siblings prior to labor market entry were driving the results, then test scores or schooling should exhibit the same asymmetric relationships as the log wage.²⁶ The first two columns of Table 7 display the impact of an individual's AFQT score on the schooling of a younger or an older sibling. Although there is some weak evidence of differences in the impact of a younger compared to an older sibling's AFQT score, the observed asymmetries have the opposite direction from those in Table 3.²⁷

A potential issue with the estimates in the first two columns of Table 7 is that AFQT scores might be an endogenous function of schooling. To address this issue, I use height instead of the AFQT score as a measure of cognitive ability.²⁸ As noted in Case and Paxson (2008), adult height is partly determined by nutritional conditions in early childhood and is positively associated with intellectual ability. The use of heights instead of AFQT scores makes it less likely that education would alter the ability measures used as regressors. The last four columns of Table 7 display the impact of an individual's height on the schooling and AFQT score of a younger or an older sibling. In all cases, the coefficient on an older sibling's height is insignificantly smaller than the coefficient on a younger sibling's height.

6 Implications for Antidiscrimination Policy

The empirical results suggest that social effects in employer learning influence wage setting. Such interactions can alter the equity and efficiency properties of labor markets. Moreover, there is potentially a role for government intervention that reduces economic inefficiencies or social inequalities. In order to address this issue formally, the online appendix presents a simple framework that demonstrates how social learning affects employment outcomes and how public policy changes the market equilibrium.

²⁶The online appendix reports additional tests of the hypothesized symmetric relationship among the human capital measures of siblings. To estimate equation (23) in the appendix, a person's AFQT score is regressed on one's own and a sibling's schooling. Furthermore, equation (26) in the appendix suggests regressing a person's height on one's own and a sibling's AFQT scores and schooling, where height is treated as an indicator of ability. Separate sets of coefficients are computed for younger and older siblings. There are no significant differences between the impacts of younger and older siblings on AFQT scores and heights.

²⁷As a further test for sibling effects on non-wage outcomes, the online appendix reports the impacts of one's own and a sibling's AFQT scores on the probabilities of marriage, children, disability, and incarceration. The estimated impacts of an older and a younger sibling's AFQT scores are small in size and similar to each other.

²⁸Information on height is available in the NLSY79 for the following survey years: 1982, 1983, 1985, 2006, 2008.

Here the lessons of the analysis are summarized, and implications for antidiscrimination programs are discussed.

An important question about nepotism is its lawfulness. Donohue (2007) outlines the main features of antidiscrimination regulation in the United States. Title VII of the Civil Rights Act of 1964 prohibits discrimination in employment on the basis of race, gender, and other protected classes. Special treatment of relatives in hiring decisions is generally regarded as legitimate, but nepotistic employment practices can be illegal if they significantly disadvantage members of a particular race or gender. This principle is known as the disparate impact doctrine, which outlaws policies that seem neutral on their face but have a discriminatory effect in practice. Nevertheless, the courts may permit employment processes that generate a disparate impact but are justifiable as a business necessity.²⁹

Consider first the case where employers follow antidiscrimination law and so do not statistically discriminate based on race. Assume, however, that information about the performance of an older relative is used when forecasting the productivity and formulating the wage of a younger relative. This practice can depress employment among youth from a disadvantaged group. In particular, an older relative from a group with lower mean productivity would perform poorly on average, generating adverse beliefs about the ability of a younger relative. Consequently, the younger relative would have worse earnings opportunities, increasing the likelihood of withdrawal from the labor force. This implication is consistent with a number of empirical studies finding that job search through personal contacts is less effective for minorities. Holzer (1987) finds that personal contacts produce fewer employment offers for blacks, and Korenman and Turner (1996) note that the wage gains from employee referrals are smaller for blacks.³⁰

Thus, nepotism may aggravate group inequalities in labor market outcomes. A government policy for equalizing employment might be a subsidy to employers for hiring younger relatives from a disadvantaged group. Examples of employer subsidies in the United States include the New Jobs Tax Credit, which was implemented between 1977 and 1978, and the Targeted Jobs Tax Credit, which lasted from 1979 to 1985 and was reinstated in 1996 in the form of the Work Opportunity Tax Credit.³¹ The former program was not limited to a specific demographic group, whereas eligibility for the latter was restricted to stigmatized populations like welfare recipients, convicted felons, and the disabled. Alternatively, policymakers might increase employment by providing an in-work benefit to members of a disadvantaged group. The main example of such a policy in the United States is the Earned Income Tax Credit, which supplements the wages of the working poor with children.³²

Suppose now that firms statistically discriminate based on group affiliation such as race, ethnicity,

²⁹For example, the Supreme Court invalidated the use of a general intelligence test for job assignment in the 1971 case of *Griggs v. Duke Power*, but the legality of a civil service exam for promotion decisions was upheld by the Supreme Court in the 2013 case of *Ricci v. DeStefano*.

³⁰Similarly, Battu et al. (2011) observe that social networks do not improve labor market outcomes among nonwhite immigrants to the United Kingdom.

³¹See Perloff and Wachter (1979) and Bishop and Montgomery (1993) for economic assessments of the New Jobs Tax Credit and the Targeted Jobs Tax Credit, respectively.

³²See Eissa and Liebman (1996) for an evaluation of the labor supply response to the Earned Income Tax Credit.

or religion. In addition, employers observe the performance of an older relative that works and can incorporate this information when forming beliefs about a younger relative. Under these assumptions, labor force participation by an older relative might be suboptimally low in a market equilibrium. By working for a firm, an older relative enables market participants to learn her productivity, and this knowledge can be useful in efficiently assigning a younger relative to an economic activity. For example, if an older relative from a disadvantaged group works and performs well, then employers may be more willing to hire a younger relative despite negative group stereotypes. Market failure may occur because an older relative ignores this positive externality of employment.

Hence, the competitive equilibrium with social learning might exhibit inadequate experimentation. This problem frequently arises in models of observational learning and strategic experimentation, where information generated by the actions of each agent serves as a public good that improves the quality of decisions undertaken by other agents. Specifically, Bolton and Harris (1999) demonstrate how a free-riding effect in a two-armed bandit problem leads to an underprovision of information, and Smith and Sørensen (2000) study incorrect herding and confounding outcomes in a model of sequential learning. In the current setting, efficiency might be restored through government interventions that increase employment among older relatives. Possible policies include hiring subsidies to firms as well as in-work benefits to employees.

7 Conclusion

This paper has implemented a test for statistical nepotism in the labor market. The empirical findings provided strong support for social effects among siblings in employer learning. This phenomenon may have important welfare implications. A possible area for extending the investigation would be to examine the schooling decisions or occupational choices of siblings. An analysis of schooling decisions might be interesting because sibling effects in employer learning could potentially contribute to the negative correlation between birth order and educational attainment documented by Behrman and Taubman (1986) and Black et al. (2005). Specifically, if an older sibling's performance provides information to employers about a younger sibling's ability, then the signaling returns to schooling may be lower for younger than for older siblings, leading younger siblings to invest less in education than their older counterparts. Alternatively, an analysis of occupational choices among siblings could help determine whether social learning can increase the efficiency of labor markets by improving the quality of information available about a worker's suitability for a given type of job.

The test for social learning is predicated on the assumption of symmetry between siblings in human capital formation. Although social effects in employer learning provide the most compelling explanation for the empirical findings, a variety of phenomena can generate differences between siblings in the structure of skills. For example, siblings may serve as role models or share human capital with each other. The characteristics of siblings might also influence resource allocation within families and parental investments in children. Such effects could in principle be asymmetric. A direction for future

research would be to extend the theoretical framework and econometric methodology in this paper to accommodate such mechanisms. One can thereby distinguish between different forms of peer and family interactions that affect labor market outcomes. Moreover, the importance of each potential channel might be quantitatively assessed.

References

- Altonji, Joseph G. and Charles R. Pierret**, “Employer Learning and Statistical Discrimination,” *Quarterly Journal of Economics*, 2001, 116 (1), 313–350.
- Autor, David H. and David Scarborough**, “Does Job Testing Harm Minority Workers? Evidence from Retail Establishments,” *Quarterly Journal of Economics*, 2008, 123 (1), 219–277.
- Battu, Harminder, Paul Seaman, and Yves Zenou**, “Job Contact Networks and the Ethnic Minorities,” *Labour Economics*, 2011, 18 (1), 48–56.
- Bayer, Patrick, Stephen L. Ross, and Giorgio Topa**, “Place of Work and Place of Residence: Informal Hiring Networks and Labor Market Outcomes,” *Journal of Political Economy*, 2008, 116 (6), 1150–1196.
- Beaman, Lori A.**, “Social Networks and the Dynamics of Labour Market Outcomes: Evidence from Refugees Resettled in the U.S.,” *Review of Economic Studies*, 2012, 79 (1), 128–161.
- **and Jeremy Magruder**, “Who Gets the Job Referral? Evidence from a Social Networks Experiment,” *American Economic Review*, 2012, 102 (7), 3574–3593.
- Becker, Gary S.**, *The Economics of Discrimination*, 2nd ed., University of Chicago Press, 1971.
- Behrman, Jere R. and Paul Taubman**, “Birth Order, Schooling, and Earnings,” *Journal of Labor Economics*, 1986, 4 (3), 121–145.
- Bishop, John H. and Mark Montgomery**, “Does the Targeted Jobs Tax Credit Create Jobs at Subsidized Firms?,” *Industrial Relations*, 1993, 32 (3), 289–306.
- Black, Sandra E., Paul Devereux, and Kjell G. Salvanes**, “The More the Merrier? The Effect of Family Size and Birth Order on Children’s Education,” *Quarterly Journal of Economics*, 2005, 120 (2), 669–700.
- Bolton, Patrick and Christopher Harris**, “Strategic Experimentation,” *Econometrica*, 1999, 67 (2), 349–374.
- Bramoullé, Yann and Gilles Saint-Paul**, “Social Networks and Labor Market Transitions,” *Labour Economics*, 2010, 17 (1), 188–195.
- Burns, Justine, Susan Godlonton, and Malcolm Keswell**, “Social Networks, Employment and Worker Discouragement: Evidence from South Africa,” *Labour Economics*, 2010, 17 (2), 336–344.
- Butcher, Kristin F. and Anne C. Case**, “The Effect of Sibling Sex Composition on Women’s Education and Earnings,” *Quarterly Journal of Economics*, 1994, 109 (3), 531–563.
- Calvò-Armengol, Antoni and Matthew O. Jackson**, “The Effects of Social Networks on Employment and Inequality,” *American Economic Review*, 2004, 94 (3), 426–454.
- Case, Anne C. and Christina Paxson**, “Height, Health, and Cognitive Function at Older Ages,” *American Economic Review: Papers and Proceedings*, 2008, 98 (2), 463–467.

- **and Lawrence F. Katz**, “The Company You Keep: The Effects of Family and Neighborhood on Disadvantaged Youth,” Working Paper, Harvard University 1991.
- Cingano, Federico and Alfonso Rosolia**, “People I Know: Job Search and Social Networks,” *Journal of Labor Economics*, 2012, 30 (2), 291–332.
- Combes, Pierre-Philippe, Laurent Linnemer, and Michael Visser**, “Publish or Peer-Rich? The Role of Skills and Networks in Hiring Economics Professors,” *Labour Economics*, 2008, 15 (3), 423–441.
- Cunha, Flavio and James Heckman**, “The Technology of Skill Formation,” *American Economic Review: Papers and Proceedings*, 2007, 97 (2), 31–47.
- DeGroot, Morris H.**, *Optimal Statistical Decisions*, McGraw-Hill, 1970.
- Donohue, John J.**, “Antidiscrimination Law,” in A. Mitchell Polinsky and Steven Shavell, eds., *Handbook of Law and Economics*, Vol. 2, North Holland, 2007, pp. 1387–1472.
- Dustmann, Christian, Albrecht Glitz, and Uta Schönberg**, “Referral-Based Job Search Networks,” Working Paper, CReAm 2011.
- Eissa, Nada and Jeffrey B. Liebman**, “Labor Supply Response to the Earned Income Tax Credit,” *Quarterly Journal of Economics*, 1996, 111 (2), 605–637.
- Farber, Henry S. and Robert Gibbons**, “Learning and Wage Dynamics,” *Quarterly Journal of Economics*, 1996, 111 (4), 1007–1047.
- Gevrek, Deniz and Z. Eylem Gevrek**, “Nepotism, Incentives, and the Academic Success of College Students,” *Labour Economics*, 2010, 17 (3), 581–591.
- Granovetter, Mark S.**, *Getting a Job: A Study of Contacts and Careers*, Harvard University Press, 1974.
- Griliches, Zvi**, “Estimating the Returns to Schooling: Some Econometric Problems,” *Econometrica*, 1977, 45 (1), 1–22.
- , “Sibling Models and Data in Economics: Beginnings of a Survey,” *Journal of Political Economy*, 1979, 87 (5), S37–S64.
- Hellerstein, Judith K. and Melinda S. Morrill**, “Dads and Daughters: The Changing Impact of Fathers on Women’s Occupational Choices,” *Journal of Human Resources*, 2011, 46 (2), 333–372.
- , **Melissa McInerney, and David Neumark**, “Neighbors and Co-Workers: The Importance of Residential Labor Market Networks,” *Journal of Labor Economics*, 2011, 29 (4), 659–695.
- Holzer, Harry J.**, “Informal Job Search and Black Youth Unemployment,” *American Economic Review*, 1987, 77 (3), 446–452.
- Ioannides, Yannis M. and Linda D. Loury**, “Job Information Networks, Neighborhood Effects, and Inequality,” *Journal of Economic Literature*, 2004, 42 (4), 1056–1093.
- Korenman, Sanders and Susan C. Turner**, “Employment Contacts and Minority-White Wage Differences,” *Industrial Relations*, 1996, 35 (1), 106–122.
- Kramarz, Francis and Oskar Nördstrom Skans**, “When Strong Ties are Strong: Networks and Youth Labor Market Entry,” Working Paper, CREST 2013.
- Kugler, Adriana D.**, “Employee Referrals and Efficiency Wages,” *Labour Economics*, 2003, 10 (5), 531–556.
- Lam, David and Robert F. Schoeni**, “Effects of Family Background on Earnings and Returns to

- Schooling: Evidence from Brazil,” *Journal of Political Economy*, 1993, 101 (4), 710–740.
- Lange, Fabian**, “The Speed of Employer Learning,” *Journal of Labor Economics*, 2007, 25 (1), 1–35.
- Manski, Charles F.**, “Identification of Endogenous Social Effects: The Reflection Problem,” *Review of Economic Studies*, 1993, 60 (3), 531–542.
- Montgomery, James D.**, “Social Networks and Labor-Market Outcomes: Toward an Economic Analysis,” *American Economic Review*, 1991, 81 (5), 1408–1418.
- Mortensen, Dale T. and Tara Vishwanath**, “Personal Contacts and Earnings: It Is Who You Know!,” *Labour Economics*, 1994, 1 (2), 187–201.
- Nakajima, Ryo, Ryuichi Tamura, and Nobuyuki Hanaki**, “The Effect of Collaboration Network on Inventors’ Job Match, Productivity and Tenure,” *Labour Economics*, 2010, 17 (4), 723–734.
- Perloff, Jeffrey M. and Michael L. Wachter**, “The New Jobs Tax Credit: An Evaluation of the 1977-78 Wage Subsidy Program,” *American Economic Review*, 1979, 69 (2), 173–179.
- Pinkston, Joshua C.**, “Screening Discrimination and the Determinants of Wages,” *Labour Economics*, 2003, 10 (6), 643–658.
- , “How Much Do Employers Learn from Referrals?,” *Industrial Relations*, 2012, 51 (2), 317–341.
- Rees, Albert E.**, “Information Networks in Labor Markets,” *American Economic Review*, 1966, 56 (1/2), 559–566.
- Simon, Curtis J. and John T. Warner**, “Matchmaker, Matchmaker: The Effect of Old Boy Networks on Job Match Quality, Earnings, and Tenure,” *Journal of Labor Economics*, 1992, 10 (3), 306–330.
- Smith, Lones and Peter Sørensen**, “Pathological Outcomes of Observational Learning,” *Econometrica*, 2000, 68 (2), 371–398.
- Wang, Shing-Yi**, “Marriage Networks, Nepotism and Labor Market Outcomes in China,” *American Economic Journal: Applied Economics*, 2013, 5 (3), 91–112.
- Zaharieva, Anna**, “Social Welfare and Wage Inequality in Search Equilibrium with Personal Contacts,” *Labour Economics*, 2013, 23, 107–121.
- Zajonc, R. B.**, “Family Configuration and Intelligence,” *Science*, 1976, 192 (4236), 227–236.

Table 1: Summary Statistics for Sibling Pairs in Labor Market

	<u>Both Siblings</u>	<u>Older Sibling</u>	<u>Younger Sibling</u>
<u>Labor Market</u>			
Real Hourly Wage			
Mean (S.D.)	7.838 (5.336)	8.101 (5.565)	7.574 (5.082)
Inter-Sib. Corr.	0.3556	—	—
<u>Human Capital</u>			
Years of Schooling			
Mean (S.D.)	12.89 (2.40)	12.95 (2.41)	12.83 (2.38)
Inter-Sib. Corr.	0.5174	—	—
Standardized AFQT			
Mean (S.D.)	-0.0566 (1.0019)	-0.0909 (1.0150)	-0.0224 (0.9875)
Inter-Sib. Corr.	0.6586	—	—
<u>Basic Demographics</u>			
Pct. Nonwhite	49.93	—	—
Pct. Female	42.43	42.30	42.55
Age			
Mean (S.D.)	31.07 (7.34)	32.34 (7.25)	29.81 (7.21)
<u>Family Background</u>			
Birth Order			
Mean (S.D.)	3.526 (2.231)	2.849 (2.075)	4.203 (2.175)
Sibship Size			
Mean (S.D.)	5.732 (2.657)	—	—
<u>Sample Size</u>			
No. Families	1993	—	—
No. Individuals	4726	2695	2707
No. Sibling Pairs	7074	—	—
No. Observations	54474	—	—

Note: The hourly wage is deflated using the CPI with 1982-1984 as the base period. The reference group for standardizing the AFQT score is all respondents in the NLSY79 having the same year of birth. Observations with missing data on a given variable are omitted when calculating the summary statistics for that variable.

Table 2: Relationship of Sibship Size and Birth Order to Probability of Sibling Helping Respondent Obtain Most Recent Job

	<u>Receive Help from Sibling</u>				<u>Also Have Same Employer</u>			
Sibship Size	0.0087 (0.0014)	—	0.0019 (0.0019)	0.0014 (0.0020)	0.0080 (0.0013)	—	0.0032 (0.0017)	0.0030 (0.0019)
Birth Order	—	0.0133 (0.0017)	0.0117 (0.0024)	0.0101 (0.0028)	—	0.0109 (0.0016)	0.0081 (0.0022)	0.0071 (0.0026)
Demographic Controls	No	No	No	Yes	No	No	No	Yes
Families	4303	4303	4303	4303	4303	4303	4303	4303
Individuals	4973	4973	4973	4973	4973	4973	4973	4973

Note: Huber-White standard errors, clustered at the family level, are reported in parentheses. All specifications control for gender and birth year fixed effects. Demographic controls include indicators for race, urban location, region of residence, parental education, parental age, and dummies for missing data on a given variable. The estimates are based on responses to the questions on job search methods in the 1982 round of the NLSY79. The sample consists of respondents who have left school for the first time and are working at a full-time job when surveyed in 1982. Individuals are excluded if they are an only child or have missing data on birth order or sibship size. During the 1982 wave of the survey, not all individuals were participating in the labor market, and so many families contain only one individual with answers to the questions about job search.

Table 3: Impact of Own AFQT and AFQT of Younger or Older Sibling in Labor Market on Log Wage

Older Sibling's AFQT \times Younger Sibling	0.0304 (0.0107)	0.0243 (0.0099)	0.0270 (0.0118)	0.0222 (0.0111)
Younger Sibling's AFQT \times Older Sibling	0.0059 (0.0102)	-0.0022 (0.0100)	-0.0142 (0.0111)	-0.0184 (0.0110)
Own AFQT \times Younger Sibling	0.0997 (0.0115)	0.0939 (0.0114)	0.1001 (0.0115)	0.0942 (0.0114)
Own AFQT \times Older Sibling	0.1600 (0.0129)	0.1495 (0.0124)	0.1650 (0.0128)	0.1550 (0.0123)
Own Schooling \times Younger Sibling	0.0524 (0.0042)	0.0498 (0.0042)	0.0516 (0.0044)	0.0493 (0.0043)
Own Schooling \times Older Sibling	0.0468 (0.0042)	0.0435 (0.0044)	0.0410 (0.0043)	0.0390 (0.0046)
Older Sibling's Schooling \times Younger Sibling	—	—	0.0025 (0.0039)	0.0017 (0.0039)
Younger Sibling's Schooling \times Older Sibling	—	—	0.0161 (0.0042)	0.0143 (0.0041)
Family Background Controls	No	Yes	No	Yes
Test for equality of ratios of own AFQT impact to sibling AFQT impact (p-value)	0.0471	0.0359	0.0083	0.0092
Families / Individuals / Sibling Pairs / Observations	1993 / 4726 / 7074 / 54474			

Note: Huber-White standard errors, clustered at the family level, are reported in parentheses. All specifications control for the race, gender, region of residence, and urban location of the members of each sibling pair. Included also are indicators for missing data on a given variable, a third-order bivariate polynomial in the ages of the two siblings, and a quartic time trend. Family background controls are indicator variables for sibship size, parental education, parental age, and each of the two siblings' birth orders. The coefficients on all control variables, except for the time trend, are estimated separately based on whether the older or the younger sibling's log wage is used as the dependent variable for a given pair. The main estimation sample described in the text is used. The p-values from the delta method are reported for the Wald test of the null hypothesis that the coefficient on (Younger Sibling's AFQT \times Older Sibling) times the coefficient on (Own AFQT \times Younger Sibling) is equal to the coefficient on (Older Sibling's AFQT \times Younger Sibling) times the coefficient on (Own AFQT \times Older Sibling).

Table 4: Impact on Log Wage of Own AFQT and AFQT of Younger or Older Sibling Residing in Different Geographic Region When First Living on Own

Older Sibling's AFQT \times Younger Sibling	-0.0141 (0.0379)	-0.0254 (0.0301)	-0.0371 (0.0418)	-0.0503 (0.0335)
Younger Sibling's AFQT \times Older Sibling	-0.0120 (0.0280)	-0.0296 (0.0287)	-0.0280 (0.0296)	-0.0373 (0.0310)
Own AFQT \times Younger Sibling	0.1346 (0.0391)	0.1237 (0.0349)	0.1379 (0.0388)	0.1298 (0.0347)
Own AFQT \times Older Sibling	0.1468 (0.0380)	0.1378 (0.0339)	0.1479 (0.0381)	0.1384 (0.0338)
Own Schooling	Yes	Yes	Yes	Yes
Sibling's Schooling	No	No	Yes	Yes
Family Background Controls	No	Yes	No	Yes
Test for equality of ratios of own AFQT impact to sibling AFQT impact (p-value)	0.9390	0.9747	0.7785	0.6727
Families / Individuals / Sibling Pairs / Observations	271 / 641 / 746 / 5698			

Note: Huber-White standard errors, clustered at the family level, are reported in parentheses. All specifications control for the race, gender, region of residence, and urban location of the members of each sibling pair. Included also are indicators for missing data on a given variable, a third-order bivariate polynomial in the ages of the two siblings, and a quartic time trend. Family background controls are indicator variables for sibship size, parental education, parental age, and each of the two siblings' birth orders. The coefficients on all control variables, except for the time trend, are estimated separately based on whether the older or the younger sibling's log wage is used as the dependent variable for a given pair. The dataset is constructed from the main estimation sample by identifying each individual's region of residence in the first survey year in which she is living in her own dwelling unit. The estimates are based on those pairs of siblings who each reside in a different region of the United States when first living on one's own. The p-values from the delta method are reported for the Wald test of the null hypothesis that the coefficient on (Younger Sibling's AFQT \times Older Sibling) times the coefficient on (Own AFQT \times Younger Sibling) is equal to the coefficient on (Older Sibling's AFQT \times Younger Sibling) times the coefficient on (Own AFQT \times Older Sibling).

Table 5: Impact on Log Wage of Own AFQT and AFQT of Younger or Older Sibling Working in Same or Different Occupation or Industry

	<u>Currently Same</u> <u>Occupation</u>		<u>Currently Same</u> <u>Industry</u>		<u>Either or Both</u>	
Older Sibling's AFQT \times Younger Sibling	0.0919 (0.0356)	0.0862 (0.0335)	0.0972 (0.0257)	0.0901 (0.0247)	0.0992 (0.0248)	0.0915 (0.0230)
Younger Sibling's AFQT \times Older Sibling	-0.0242 (0.0332)	-0.0184 (0.0317)	-0.0304 (0.0269)	-0.0094 (0.0258)	-0.0250 (0.0247)	-0.0124 (0.0233)
Own AFQT \times Younger Sibling	0.0548 (0.0334)	0.0461 (0.0345)	0.0640 (0.0281)	0.0740 (0.0270)	0.0529 (0.0253)	0.0539 (0.0254)
Own AFQT \times Older Sibling	0.2127 (0.0271)	0.1953 (0.0275)	0.1514 (0.0279)	0.1227 (0.0270)	0.1731 (0.0234)	0.1491 (0.0234)
Own and Sibling's Schooling	Yes	Yes	Yes	Yes	Yes	Yes
Family Background Controls	No	Yes	No	Yes	No	Yes
Test for equality of ratios of own AFQT impact to sibling AFQT impact (p-value)	0.0230	0.0265	0.0098	0.0348	0.0029	0.0066
Families	486	486	587	587	746	746
Individuals	1060	1060	1290	1290	1674	1674
Sibling Pairs	1202	1202	1476	1476	1968	1968
Observations	2424	2424	4040	4040	5364	5364
	<u>Always Different</u> <u>Occupation</u>		<u>Always Different</u> <u>Industry</u>		<u>Both</u>	
Older Sibling's AFQT \times Younger Sibling	-0.0075 (0.0160)	-0.0122 (0.0156)	0.0130 (0.0174)	0.0094 (0.0162)	-0.0044 (0.0204)	-0.0071 (0.0199)
Younger Sibling's AFQT \times Older Sibling	-0.0114 (0.0164)	-0.0157 (0.0166)	-0.0085 (0.0174)	-0.0188 (0.0172)	-0.0004 (0.0217)	-0.0055 (0.0214)
Own AFQT \times Younger Sibling	0.1153 (0.0177)	0.1107 (0.0174)	0.0951 (0.0162)	0.0860 (0.0163)	0.0952 (0.0223)	0.0879 (0.0226)
Own AFQT \times Older Sibling	0.1271 (0.0185)	0.1198 (0.0182)	0.1652 (0.0215)	0.1549 (0.0206)	0.1364 (0.0249)	0.1305 (0.0239)
Own and Sibling's Schooling	Yes	Yes	Yes	Yes	Yes	Yes
Family Background Controls	No	Yes	No	Yes	No	Yes
Test for equality of ratios of own AFQT impact to sibling AFQT impact (p-value)	0.8893	0.9105	0.3781	0.2929	0.8729	0.8881
Families	947	947	902	902	609	609
Individuals	2220	2220	2117	2117	1390	1390
Sibling Pairs	2692	2692	2576	2576	1606	1606
Observations	17512	17512	17732	17732	10306	10306

Note: Huber-White standard errors, clustered at the family level, are reported in parentheses. All specifications control for the race, gender, region of residence, and urban location of the members of each sibling pair. Included also are indicators for missing data on a given variable, a third-order bivariate polynomial in the ages of the two siblings, and a quartic time trend. Family background controls are indicator variables for sibship size, parental education, parental age, and each of the two siblings' birth orders. The coefficients on all control variables, except for the time trend, are estimated separately based on whether the older or the younger sibling's log wage is used as the dependent variable for a given pair. A pair of siblings is labeled as currently having the same occupation (industry) if they both belong to the same occupation (industry) in the relevant survey year. Two siblings are said to always be in different occupations (industries) if the set of occupations (industries) reported by one sibling is disjoint from the set of occupations (industries) reported by the other sibling over the entire course of the survey. The 2000 Census 3-digit occupation and industry codes are used to classify observations on sibling pairs. Between the 1979 and 2000 rounds of the NLSY79, the occupation and industry of each job were originally recorded as 1970 Census 3-digit codes. These fields are converted to 2000 Census 3-digit codes based on the crosswalks available from the US Census Bureau. The p-values from the delta method are reported for the Wald test of the null hypothesis that the coefficient on (Younger Sibling's AFQT \times Older Sibling) times the coefficient on (Own AFQT \times Younger Sibling) is equal to the coefficient on (Older Sibling's AFQT \times Younger Sibling) times the coefficient on (Own AFQT \times Older Sibling).

Table 6: Impact on Log Wage for New Hires and Senior Employees of Own AFQT and AFQT of Sibling Working in Same or Different Occupation or Industry

	Same Occupation or Industry		Different Occupation and Industry	
Sibling's AFQT \times New Hire	0.1082 (0.0307)	0.1312 (0.0333)	0.0112 (0.0135)	0.0111 (0.0135)
Sibling's AFQT \times Senior Employee	0.0324 (0.0234)	0.0254 (0.0231)	0.0099 (0.0106)	0.0020 (0.0102)
Own AFQT \times New Hire	0.0537 (0.0366)	0.0535 (0.0390)	0.0962 (0.0146)	0.0929 (0.0141)
Own AFQT \times Senior Employee	0.1037 (0.0200)	0.0851 (0.0216)	0.1309 (0.0111)	0.1225 (0.0112)
Own and Sibling's Schooling	Yes	Yes	Yes	Yes
Family Background Controls	No	Yes	No	Yes
Test for equality of ratios of own AFQT impact to sibling AFQT impact (p-value)	0.0363	0.0329	0.7877	0.5110
Families	619	619	1826	1826
Individuals	1107	1107	3904	3904
Sibling Pairs	1276	1276	5657	5657
Observations	3044	3044	23998	23998

Note: Huber-White standard errors, clustered at the family level, are reported in parentheses. All specifications control for the race, gender, region of residence, and urban location of the members of each sibling pair. Included also are indicators for missing data on a given variable, a third-order bivariate polynomial in the ages of the two siblings, and a quartic time trend. Family background controls are indicator variables for sibship size, parental education, parental age, and each of the two siblings' birth orders. The coefficients on all control variables, except for the time trend, are estimated separately based on whether the dependent variable is the log wage observation for a new hire or a senior employee. The dataset used here derives from the main estimation sample. An individual with under three months of tenure at her current employer is labeled as a new hire. A person with over three years of tenure on the job is classified as a senior employee. The 2000 Census 3-digit occupation and industry codes are used to categorize observations on sibling pairs. Between the 1979 and 2000 rounds of the NLSY79, the occupation and industry of each job were originally recorded as 1970 Census 3-digit codes. These fields are converted to 2000 Census 3-digit codes based on the crosswalks available from the US Census Bureau. The p-values from the delta method are reported for the Wald test of the null hypothesis that the coefficient on (Sibling's AFQT \times Senior Employee) times the coefficient on (Own AFQT \times New Hire) is equal to the coefficient on (Sibling's AFQT \times New Hire) times the coefficient on (Own AFQT \times Senior Employee).

Table 7: Relationship of Own AFQT and Schooling to AFQT and Height of Younger or Older Sibling

	Schooling		AFQT			
Older Sibling's AFQT \times Younger Sibling	0.3117 (0.0416)	0.1334 (0.0434)	—	—	—	—
Younger Sibling's AFQT \times Older Sibling	0.3859 (0.0385)	0.2132 (0.0410)	—	—	—	—
Own AFQT	1.3927 (0.0395)	1.1932 (0.0420)	—	—	—	—
Older Sibling's Height \times Younger Sibling	—	—	0.0424 (0.0140)	0.0094 (0.0129)	0.0150 (0.0050)	0.0030 (0.0044)
Younger Sibling's Height \times Older Sibling	—	—	0.0549 (0.0139)	0.0214 (0.0123)	0.0203 (0.0053)	0.0084 (0.0046)
Own Height	—	—	0.0772 (0.0129)	0.0386 (0.0119)	0.0281 (0.0048)	0.0142 (0.0043)
Family Background Controls	No	Yes	No	Yes	No	Yes
Test for equality between sibling AFQT or height coefficients (p-value)	0.0726	0.0548	0.3594	0.3273	0.2851	0.2196
Families / Individuals / Sibling Pairs	1993 / 4726 / 7074					

Note: Huber-White standard errors, clustered at the family level, are reported in parentheses. All specifications control for the race, gender, region of residence, and urban location of the members of each sibling pair. Included also are indicators for missing data on a given variable and fixed effects for each of the two siblings' years of birth. Family background controls are indicator variables for sibship size, parental education, parental age, and each of the two siblings' birth orders. All specifications include an indicator for whether the respondent is the older or the younger sibling in a given pair. The dataset contains the first observation on every sibling pair in the main estimation sample. However, the second four columns exclude sibling pairs in which either member is missing information on height. The first two columns report p-values for the Wald test of the restriction that the coefficient on (Older Sibling's AFQT \times Younger Sibling) is equal to the coefficient on (Younger Sibling's AFQT \times Older Sibling). The second four columns report p-values for the Wald test of the restriction that the coefficient on (Older Sibling's Height \times Younger Sibling) is equal to the coefficient on (Younger Sibling's Height \times Older Sibling).

Appendix

This appendix contains the proof of proposition 2. The following notation is used. Let \hat{X} and \hat{Y} be random variables. The expectation and variance of \hat{X} are respectively denoted by $\mathbb{E}(\hat{X})$ and $\mathbb{V}(\hat{X})$. The conditional expectation and conditional variance of \hat{X} given \hat{Y} are respectively denoted by $\mathbb{E}(\hat{X}|\hat{Y})$ and $\mathbb{V}(\hat{X}|\hat{Y})$. The notation for the covariance between \hat{X} and \hat{Y} is $\mathbb{C}(\hat{X}, \hat{Y})$. If \hat{X} and \hat{Y} are random vectors, then $\mathbb{V}(\hat{X})$ represents the variance-covariance matrix of \hat{X} , and $\mathbb{C}(\hat{X}, \hat{Y})$ represents the cross-covariance matrix between \hat{X} and \hat{Y} .

I Lemmata for Proof of Proposition

In order to derive the empirical implications of the social learning model, it is first necessary to analyze the coefficient obtained from a hypothetical regression of the siblings' abilities a_1, a_2 on their test scores z_1, z_2 after controlling for their schooling s_1, s_2 . Let σ_v^2 be the variance of the variable v_i and ρ_v be the correlation between v_1 and v_2 . The analysis proceeds in two steps.

The first step is to calculate the component of each sibling's test score that is orthogonal to her own and her sibling's schooling. The result below characterizes the problem of predicting the test scores z_1, z_2 from a regression on the schooling levels s_1, s_2 .

Lemma 1 *The regression coefficient of $(z_1, z_2)'$ on $(s_1, s_2)'$ is given by:*

$$\mathbb{C} \left[\begin{pmatrix} z_1 \\ z_2 \end{pmatrix}, \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} \right] \mathbb{V} \left[\begin{pmatrix} s_1 \\ s_2 \end{pmatrix} \right]^{-1} = \theta_s \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \theta_a \frac{\gamma \sigma_a^2}{\sigma_s^2 (1 - \rho_s^2)} \begin{pmatrix} 1 - \rho_a \rho_s & \rho_a - \rho_s \\ \rho_a - \rho_s & 1 - \rho_a \rho_s \end{pmatrix}. \quad (23)$$

Proof The conditional expectation of $(z_1, z_2)'$ given $(s_1, s_2)'$ is:

$$\mathbb{E} \left[\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \mid \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} \right] = \begin{pmatrix} \mu_{z1} \\ \mu_{z2} \end{pmatrix} + \mathbb{C} \left[\begin{pmatrix} z_1 \\ z_2 \end{pmatrix}, \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} \right] \mathbb{V} \left[\begin{pmatrix} s_1 \\ s_2 \end{pmatrix} \right]^{-1} \left[\begin{pmatrix} s_1 \\ s_2 \end{pmatrix} - \begin{pmatrix} \mu_{s1} \\ \mu_{s2} \end{pmatrix} \right], \quad (24)$$

where the regression coefficient is given by:

$$\mathbb{C} \left[\begin{pmatrix} z_1 \\ z_2 \end{pmatrix}, \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} \right] \mathbb{V} \left[\begin{pmatrix} s_1 \\ s_2 \end{pmatrix} \right]^{-1} = \left[\theta_s \sigma_s^2 \begin{pmatrix} 1 & \rho_s \\ \rho_s & 1 \end{pmatrix} + \theta_a \gamma \sigma_a^2 \begin{pmatrix} 1 & \rho_a \\ \rho_a & 1 \end{pmatrix} \right] \begin{pmatrix} \sigma_s^2 & \rho_s \sigma_s^2 \\ \rho_s \sigma_s^2 & \sigma_s^2 \end{pmatrix}^{-1}. \quad (25)$$

Inverting the variance matrix and rearranging terms leads to the formula in equation (23). ■

The first term on the right-hand side of equation (23) accounts for the causal effect of schooling on the test score. The second term, which arises from the relationship of schooling with ability, is a generalization of the univariate measurement error formula, where schooling is treated as ability measured with error.³³

This formula provides a method for answering one of the central questions raised by Griliches (1979) regarding the role of families in human capital formation.³⁴ In particular, one can directly test whether the sibling correlation ρ_a in ability is greater or less than the sibling correlation ρ_s in schooling, especially

³³Consider the special case where $\rho_a = \rho_s$ and $\gamma = 1$. Then, in the second term, the coefficient on one's own schooling is the parameter θ_a multiplied by the reliability ratio σ_a^2/σ_s^2 . The coefficient on a sibling's schooling is zero in this case.

³⁴An analogous formula applies if the log wage instead of the test score is used as the dependent variable, provided that the log wage can be modeled like the test score as a linear combination of ability, schooling, and an error term.

if the schooling variable itself is not measured with error. Equation (23) shows that if $(z_1, z_2)'$ is regressed on $(s_1, s_2)'$, then the coefficient on a sibling's schooling is positive if $\rho_a > \rho_s$ and negative if $\rho_a < \rho_s$. This question is relevant when interpreting family fixed-effects estimates of the returns to schooling. Depending on whether siblings have a higher or lower correlation in ability than in schooling, the within-family estimator of the schooling coefficient may either mitigate or exacerbate ability bias relative to the ordinary least squares estimate.

The second step is to characterize the relationship between siblings' abilities and test scores after partialling out the influence of schooling. Denoting $s = (s_1, s_2)'$ and $z = (z_1, z_2)'$, consider the regression coefficient of the siblings' abilities on their schooling and test scores:

$$\mathbb{C} \left[\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \begin{pmatrix} s \\ z \end{pmatrix} \right] \mathbb{V} \left[\begin{pmatrix} s \\ z \end{pmatrix} \right]^{-1} = \begin{pmatrix} \psi_o & \psi_f & \pi_o & \pi_f \\ \psi_f & \psi_o & \pi_f & \pi_o \end{pmatrix}. \quad (26)$$

In equation (26), the coefficients π_o and π_f on one's own and a sibling's test scores do not differ between the two siblings, and likewise for the coefficients ψ_o and ψ_f on one's own and a sibling's schooling. This result obtains because the skills of younger and older siblings are modeled symmetrically. The specifications of schooling and test scores are analogous for both siblings in equations (3) and (5). The impact γ of ability on schooling does not vary between siblings, nor do the impacts θ_s and θ_a of schooling and ability on test scores. Moreover, the abilities a_1 and a_2 of the two siblings have an identical variance, although the means can differ. The same holds for the error terms ϵ_1 and ϵ_2 in the schooling equation, which are independent of a_1 and a_2 , and for the error terms ω_1 and ω_2 in the test score equation, which are independent of a_1 and a_2 as well as ϵ_1 and ϵ_2 . The similar treatment of siblings is key to proposition 1, which shows that the ratio of the coefficients on test scores does not vary between siblings under individual learning. Given the underlying symmetry of the human capital measures, the asymmetries in proposition 2 can be attributed to social learning.

The result below enumerates the basic properties of the regression parameters π_o and π_f , which represent the relationship of one's ability to one's own test score and a sibling's test score.

Lemma 2 *The regression parameters π_o and π_f satisfy $\pi_o > \pi_f$, $\pi_o > 0$, and $\pi_o^2 > \pi_f^2$.*

Proof Expressing the regression coefficient in equation (23) as:

$$\mathbb{C} \left[\begin{pmatrix} z_1 \\ z_2 \end{pmatrix}, \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} \right] \mathbb{V} \left[\begin{pmatrix} s_1 \\ s_2 \end{pmatrix} \right]^{-1} = \begin{pmatrix} \theta_s + \theta_a \delta_o & \theta_a \delta_f \\ \theta_a \delta_f & \theta_s + \theta_a \delta_o \end{pmatrix}, \quad (27)$$

the component of $(z_1, z_2)'$ orthogonal to $(s_1, s_2)'$ is given by:

$$\begin{pmatrix} \hat{z}_1 \\ \hat{z}_2 \end{pmatrix} = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} - \mathbb{E} \left[\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \mid \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} \right] = \begin{pmatrix} \theta_a [a_1 - (\delta_o s_1 + \delta_f s_2)] + \omega_1 \\ \theta_a [a_2 - (\delta_f s_1 + \delta_o s_2)] + \omega_2 \end{pmatrix} - \begin{pmatrix} \theta_a [\mu_{a1} - (\delta_o \mu_{s1} + \delta_f \mu_{s2})] \\ \theta_a [\mu_{a2} - (\delta_f \mu_{s1} + \delta_o \mu_{s2})] \end{pmatrix}, \quad (28)$$

where equations (5) and (24) are used to substitute for $(z_1, z_2)'$ and $\mathbb{E}[(z_1, z_2)' | (s_1, s_2)']$, respectively. Note that the coefficient on $(z_1, z_2)'$ in a regression on $(s_1, s_2, z_1, z_2)'$ is the same as the coefficient on $(\hat{z}_1, \hat{z}_2)'$ in a regression on $(s_1, s_2, \hat{z}_1, \hat{z}_2)'$. Therefore, consider the regression of $(a_1, a_2)'$ on $(s_1, s_2, \hat{z}_1, \hat{z}_2)'$. Because $(\hat{z}_1, \hat{z}_2)'$ is uncorrelated with $(s_1, s_2)'$, the coefficient on $(\hat{z}_1, \hat{z}_2)'$ is simply given by:

$$\begin{pmatrix} \pi_o & \pi_f \\ \pi_f & \pi_o \end{pmatrix} = \mathbb{C} \left[\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \begin{pmatrix} \hat{z}_1 \\ \hat{z}_2 \end{pmatrix} \right] \mathbb{V} \left[\begin{pmatrix} \hat{z}_1 \\ \hat{z}_2 \end{pmatrix} \right]^{-1}, \quad (29)$$

where the inverse variance matrix can be expressed as:

$$\mathbb{V} \left[\begin{pmatrix} \hat{z}_1 \\ \hat{z}_2 \end{pmatrix} \right]^{-1} = [\sigma_{\hat{z}}^2(1 - \rho_{\hat{z}}^2)]^{-1} \begin{pmatrix} \sigma_{\hat{z}}^2 & -\rho_{\hat{z}}\sigma_{\hat{z}}^2 \\ -\rho_{\hat{z}}\sigma_{\hat{z}}^2 & \sigma_{\hat{z}}^2 \end{pmatrix}. \quad (30)$$

I first show that $\mathbb{C}(a_1, \hat{z}_1) = \mathbb{C}(a_2, \hat{z}_2) > 0$. Using equations (28) and (23), this covariance is given by:

$$\mathbb{C}(a_1, \hat{z}_1) = \mathbb{C}(a_2, \hat{z}_2) = \theta_a \sigma_a^2 [1 - \gamma(\delta_o + \delta_f \rho_a)] = \theta_a \sigma_a^2 \left(1 - \frac{\gamma^2 \sigma_a^2 (1 - 2\rho_a \rho_s + \rho_a^2)}{\sigma_s^2 (1 - \rho_s^2)} \right). \quad (31)$$

From equation (31), the statement $\mathbb{C}(a_1, \hat{z}_1) = \mathbb{C}(a_2, \hat{z}_2) > 0$ is equivalent to:

$$\sigma_s^2 (1 - \rho_s^2) - \gamma^2 \sigma_a^2 (1 - 2\rho_a \rho_s + \rho_a^2) > 0, \quad (32)$$

which can be expanded as:

$$k_o = (\gamma^2 \sigma_a^2 + \sigma_\epsilon^2) \left(1 - \frac{(\gamma^2 \rho_a \sigma_a^2 + \rho_\epsilon \sigma_\epsilon^2)^2}{(\gamma^2 \sigma_a^2 + \sigma_\epsilon^2)^2} \right) - \gamma^2 \sigma_a^2 \left(1 - 2\rho_a \frac{\gamma^2 \rho_a \sigma_a^2 + \rho_\epsilon \sigma_\epsilon^2}{\gamma^2 \sigma_a^2 + \sigma_\epsilon^2} + \rho_a^2 \right) > 0. \quad (33)$$

Note that $k_o = 0$ when $\sigma_\epsilon^2 = 0$. The derivative of k_o with respect to σ_ϵ^2 is:

$$\frac{\partial k_o}{\partial \sigma_\epsilon^2} = \frac{\sigma_\epsilon^2 (1 - \rho_\epsilon^2) + 2\gamma^2 \sigma_a^2 \sigma_\epsilon^2 (1 - \rho_\epsilon^2) + \gamma^4 \sigma_a^4 (1 - \rho_a^2)}{(\gamma^2 \sigma_a^2 + \sigma_\epsilon^2)^2} > 0. \quad (34)$$

It follows that $k_o > 0$ if $\sigma_\epsilon^2 > 0$ and so $\mathbb{C}(a_1, \hat{z}_1) = \mathbb{C}(a_2, \hat{z}_2) > 0$ in equation (31).

I next show that $\mathbb{C}(a_1, \hat{z}_2) = \mathbb{C}(a_2, \hat{z}_1) > 0$. Using equations (28) and (23), this covariance is given by:

$$\mathbb{C}(a_1, \hat{z}_2) = \mathbb{C}(a_2, \hat{z}_1) = \theta_a \sigma_a^2 [\rho_a - \gamma(\delta_o \rho_a + \delta_f)] = \theta_a \sigma_a^2 \left(\rho_a - \frac{\gamma^2 \sigma_a^2 [2\rho_a - \rho_s (1 + \rho_a^2)]}{\sigma_s^2 (1 - \rho_s^2)} \right). \quad (35)$$

From equation (35), the statement $\mathbb{C}(a_1, \hat{z}_2) = \mathbb{C}(a_2, \hat{z}_1) > 0$ is equivalent to:

$$\rho_a \sigma_s^2 (1 - \rho_s^2) - \gamma^2 \sigma_a^2 [2\rho_a - \rho_s (1 + \rho_a^2)] > 0, \quad (36)$$

which can be expanded as:

$$k_f = \rho_a (\gamma^2 \sigma_a^2 + \sigma_\epsilon^2) \left(1 - \frac{(\gamma^2 \rho_a \sigma_a^2 + \rho_\epsilon \sigma_\epsilon^2)^2}{(\gamma^2 \sigma_a^2 + \sigma_\epsilon^2)^2} \right) - \gamma^2 \sigma_a^2 \left(2\rho_a - (1 + \rho_a^2) \frac{\gamma^2 \rho_a \sigma_a^2 + \rho_\epsilon \sigma_\epsilon^2}{\gamma^2 \sigma_a^2 + \sigma_\epsilon^2} \right) > 0. \quad (37)$$

Note that $k_f = 0$ when $\sigma_\epsilon^2 = 0$. The derivative of k_f with respect to σ_ϵ^2 is:

$$\frac{\partial k_f}{\partial \sigma_\epsilon^2} = \frac{\gamma^4 \sigma_a^4 \rho_\epsilon (1 - \rho_a^2) + \sigma_\epsilon^2 (2\gamma^2 \sigma_a^2 + \sigma_\epsilon^2) \rho_a (1 - \rho_\epsilon^2)}{(\gamma^2 \sigma_a^2 + \sigma_\epsilon^2)^2} > 0. \quad (38)$$

It follows that $k_f > 0$ if $\sigma_\epsilon^2 > 0$ and so $\mathbb{C}(a_1, \hat{z}_2) = \mathbb{C}(a_2, \hat{z}_1) > 0$ in equation (35).

I now show that $\mathbb{C}(a_1, \hat{z}_1) > \mathbb{C}(a_1, \hat{z}_2)$. From equations (31) and (35), the statement $\mathbb{C}(a_1, \hat{z}_1) > \mathbb{C}(a_1, \hat{z}_2)$ is equivalent to:

$$1 - \frac{\gamma^2 \sigma_a^2 (1 - 2\rho_a \rho_s + \rho_a^2)}{\sigma_s^2 (1 - \rho_s^2)} > \rho_a - \frac{\gamma^2 \sigma_a^2 (2\rho_a - \rho_a^2 \rho_s - \rho_s)}{\sigma_s^2 (1 - \rho_s^2)}, \quad (39)$$

which is satisfied whenever:

$$w = \sigma_s^2(1 - \rho_s) - \gamma^2 \sigma_a^2(1 - \rho_a) = (\gamma^2 \sigma_a^2 + \sigma_\epsilon^2) \left(1 - \frac{\gamma^2 \rho_a \sigma_a^2 + \rho_\epsilon \sigma_\epsilon^2}{\gamma^2 \sigma_a^2 + \sigma_\epsilon^2} \right) - \gamma^2 \sigma_a^2(1 - \rho_a) > 0. \quad (40)$$

Note that $w = 0$ if $\sigma_\epsilon^2 = 0$. Differentiating w with respect to σ_ϵ^2 yields:

$$\frac{\partial w}{\partial \sigma_\epsilon^2} = 1 - \rho_\epsilon > 0. \quad (41)$$

Hence, $w > 0$ if $\sigma_\epsilon^2 > 0$ and so $\mathbb{C}(a_1, \hat{z}_1) > \mathbb{C}(a_1, \hat{z}_2)$.

It is now straightforward to prove the three claims in the lemma. Given the form of the inverse variance matrix in equation (30), it follows from $\mathbb{C}(a_1, \hat{z}_1) > \mathbb{C}(a_1, \hat{z}_2)$ that $\pi_o > \pi_f$, proving the first claim. From equations (29) and (30), the regression parameters π_o and π_f take the form:

$$\pi_o = \kappa_o - \rho_z \kappa_f \quad \text{and} \quad \pi_f = \kappa_f - \rho_z \kappa_o, \quad (42)$$

where $\kappa_o = \tau \mathbb{C}(a_1, \hat{z}_1) > 0$, $\kappa_f = \tau \mathbb{C}(a_1, \hat{z}_2) > 0$, and $\tau > 0$. Because it has been shown above that $\mathbb{C}(a_1, \hat{z}_1) > \mathbb{C}(a_1, \hat{z}_2) > 0$, one has $\kappa_o > \kappa_f > 0$. These inequalities imply that $\pi_o > 0$ in equation (42), proving the second claim. Finally, because $\kappa_o^2 > \kappa_f^2$ from the preceding inequalities, one has:

$$\kappa_o^2 + \rho_z^2 \kappa_f^2 > \rho_z^2 \kappa_o^2 + \kappa_f^2 \Leftrightarrow \kappa_o^2 - 2\rho_z \kappa_o \kappa_f + \rho_z^2 \kappa_f^2 > \rho_z^2 \kappa_o^2 - 2\rho_z \kappa_o \kappa_f + \kappa_f^2 \Leftrightarrow (\kappa_o - \rho_z \kappa_f)^2 > (\kappa_f - \rho_z \kappa_o)^2; \quad (43)$$

so that, $\pi_o^2 > \pi_f^2$ in equation (42), proving the third claim. \blacksquare

The three parts of lemma 2 can be stated as follows. First, one's own test score remains a stronger predictor of one's ability than a sibling's test score after controlling for one's own and a sibling's schooling. Second, the partial correlation of one's ability with one's own test score is positive. Third, the coefficient on one's own test score is larger in absolute value than the coefficient on a sibling's test score in the regression of one's ability on one's own and a sibling's test scores and schooling. These simple properties are important in deriving the empirical implications of the employer learning models.

Although lemma 2 demonstrates that one's own test score is positively related to one's ability given the other regressors, there is no analogous result for the coefficient on a sibling's test score, which can in general have either a positive or a negative partial correlation with one's ability. The reason for this ambiguity is that the test score is affected by factors other than ability that may be correlated between siblings. In other words, the sign of the coefficient π_f on a sibling's test score is the outcome of two competing effects: a positive correlation in ability a_i leads π_f to be positive, but a positive correlation in testing error ω_i leads π_f to be negative. This feature of the model is akin to the finding in lemma 1 that a sibling's schooling can have either a positive or a negative coefficient in the regression of one's test score on one's own and a sibling's schooling.

II Proof of Proposition

To prove the second part of the proposition, I first calculate the coefficient ζ_{ri} on \bar{r}_e in the conditional expectation of a_i given s_i , s_e , and \bar{r}_e in equation (18). The component of \bar{r}_e orthogonal to s_i and s_e is given by:

$$\hat{r}_e = \bar{r}_e - \mathbb{E}(\bar{r}_e | s_i, s_e) = [a_e - (\delta_o s_e + \delta_f s_i) + \bar{\eta}_e] - [\mu_{ae} - (\delta_o \mu_{se} + \delta_f \mu_{si})], \quad (44)$$

where δ_o, δ_f are defined in equation (27), and $\bar{\eta}_e$ is the sample mean of $\{\eta_{ue}\}_{u=1}^{t_e}$. Note that the coefficient on \bar{r}_e in the conditional expectation given s_i, s_e , and \bar{r}_e is the same as the coefficient on \hat{r}_e in the conditional expectation given s_i, s_e , and \hat{r}_e . Because \hat{r}_e is uncorrelated with s_i and s_e , the coefficient ζ_{ri} is equal to:

$$\zeta_{ri} = \mathbb{C}(a_i, \hat{r}_e) / \mathbb{V}(\hat{r}_e) = \sigma_a^2 [\rho_a - \gamma(\delta_o \rho_a + \delta_f)] / (\zeta^2 + t_e^{-1} \sigma_\eta^2), \quad (45)$$

where ζ^2 is defined as:

$$\zeta^2 = \mathbb{V}[a_e - (\delta_o s_e + \delta_f s_i)]. \quad (46)$$

Note that the bracketed term in the numerator of equation (45) also appears in equation (35) and was shown to be positive in the proof of lemma 2. Thus, ζ_{ri} is positive. Moreover, if $t_1 > t_2$, then $\zeta_{r1} < \zeta_{r2}$.

From equation (21), the coefficients ν_{ie}, ν_{ii} in proposition 2 can be expressed as:

$$\nu_{ii} = (1 - \xi_i) \zeta_{ri} \pi_f + \xi_i \pi_o \quad \text{and} \quad \nu_{ie} = (1 - \xi_i) \zeta_{ri} \pi_o + \xi_i \pi_f. \quad (47)$$

Thus, the statement $\nu_{12} \nu_{22} < \nu_{21} \nu_{11}$ is equivalent to:

$$[(1 - \xi_1) \zeta_{r1} \pi_o + \xi_1 \pi_f] [(1 - \xi_2) \zeta_{r2} \pi_f + \xi_2 \pi_o] < [(1 - \xi_2) \zeta_{r2} \pi_o + \xi_2 \pi_f] [(1 - \xi_1) \zeta_{r1} \pi_f + \xi_1 \pi_o], \quad (48)$$

which reduces to:

$$(1 - \xi_1) \xi_2 \zeta_{r1} \pi_o^2 + \xi_1 (1 - \xi_2) \zeta_{r2} \pi_f^2 < (1 - \xi_2) \xi_1 \zeta_{r2} \pi_o^2 + \xi_2 (1 - \xi_1) \zeta_{r1} \pi_f^2. \quad (49)$$

If $t_1 > t_2$, then $\xi_1 > \xi_2$ and $\zeta_{r1} < \zeta_{r2}$. Thus, equation (49) is satisfied if $\pi_o^2 > \pi_f^2$ holds, and lemma 2 shows that $\pi_o^2 > \pi_f^2$.