SFI THE DANISH NATIONAL CENTRE FOR SOCIAL RESEARCH

Alice Heegaard Klynge

04:2015 WORKING PAPER

Occupational Choice and the Endogenous Supply of Ability

SFI – THE DANISH NATIONAL CENTRE FOR SOCIAL RESEARCH

OCCUPATIONAL CHOICE AND THE ENDOGENOUS SUPPLY OF ABILITY

Alice Heegaard Klynge

THE DANISH NATIONAL CENTRE FOR SOCIAL RESEARCH, COPENHAGEN, DENMARK

Working Paper 04:2015

Occupational Choice and the Endogenous Supply of Ability

Alice Heegaard Klynge*

September 7, 2015

Abstract

I develop a selection model in which the individual's supply of ability is endogenous and subject to selection along with occupation. Additionally, I identify and estimate the returns to creative and innovative ability, communication ability, and reading and math ability for white-collar and blue-collar workers. The model permits a person's choice of occupation to influence his decision regarding the amount of ability to supply. The empirical results show that the individual's supply of ability should be allowed to be endogenous to correct for sample selection bias caused by occupational choice.

JEL Codes: J08, J22, J24, J31, J32, J62

Keywords: Labor economics, labor supply, human capital, selection model, labor productivity, returns to noncognitive abilities, structural estimation, endogenous regressors

* The Danish National Centre for Social Research. E-mail contact: ahk@sfi.dk. Web: www.sfi.dk/klynge. I especially thank Edward Vytlacil, Martin Browning, and Bo Honoré for insightful comments. I also thank Signe Hald Andersen, Reena Badiani, Kathrine Bløcher, Sylvain Chabe-Ferret, Arnaud Dupuy, Mette Ejrnæs, Kirill Evdokimov, Glenn W. Harrison, James J. Heckman, Anders Holm, Niels Johannesen, Mads M. Jæger, Dennis Kristensen, Fabian Lange, Aureo de Paula, Bertel Schjerning, Xiaoxia Shi, Frederik Silbye, Melissa Tartari, and Nese Yildiz for valuable comments. Additionally, I thank seminar participants at Yale University, Princeton University, Columbia University, University of Copenhagen, the Applied Microeconometrics Conference at Oxford University, the Danish Microeconometric Network Meeting, and the European Economic Association Meeting 2013 for valuable comments.

I. Introduction

An individual's choice of occupation may influence his decision regarding the amount of ability to supply. For example, a person may decide to use more of his creative and innovative ability as a professor than as a garbage collector. Neglecting the influence of choice of occupation on the extent to which an individual uses his abilities in a conventional selection model leads to a specification error in the population model and to a misspecification error in the estimated returns to abilities within an occupation.

In this paper, I develop a selection model in which the individual's supply of ability is endogenous and subject to selection along with occupation. By comparison, the Roy model and later applications of the selection model to choice of occupations or sectors have assumed that measured ability is constant across occupations or sectors (e.g., Roy 1951; Heckman and Sedlacek 1985, 1990; Willis 1986; Heckman and Honoré 1990; Cawley, Heckman, and Vytlacil 1999; Gould 2002). This assumption is sensible for innate abilities, which is the primary focus of these studies.

The selection model developed here builds on the Roy model in which each worker chooses his occupation. Additionally, the model builds on Heckman and Sedlacek (1985, 1990), Willis (1986), and Gould (2002), who formalize the Roy model and allow each ability to be useful in every occupation. I extend the previous literature with the idea that income depends on the supply of ability rather than the innate ability. How much the worker actually uses his ability determines his income rather than the level he potentially could use. Moreover, I permit individuals optimally to choose how much ability to devote toward a given occupation. As a result, an individual's choice of occupation influences his decision regarding the amount of ability to supply.

Two challenges arise in identifying and estimating returns to abilities within an occupation. The first is a selection problem because income is only observed in occupations that are chosen. The second is a partial observation problem: how much ability a worker would have supplied in each occupation he could have chosen is unobserved. The selection problem is a standard challenge in selection models and is solved by including an exclusion restriction that influences the choice of occupation but not income (e.g., Heckman 1976, 1979).

The partial observation of the supply of ability in each occupation poses a special challenge. My solution involves projecting the supply of ability in work on the supply of ability in leisure for each type of ability. For example, for each individual, the supply of creative and innovative ability in work is projected on the supply of creative and innovative ability in leisure.[1] The model yields a person's optimal supply of ability in each occupation and in leisure as a function of his innate ability. The model thus allows me to determine how much ability an individual would have used in each occupation he could have chosen.

Based on the structure of the model, I develop a new estimator of the return to ability within an occupation. The estimator is identified through a three-step procedure. The approach builds on the two-step procedure of a conventional selection model by adding a step in which the supply of ability in work is projected on the supply of ability in leisure for each type of ability. Moreover, the explanatory variables in every step are the abilities supplied in leisure, rather than the abilities supplied in work, as the abilities supplied in leisure substitute for the partially observed abilities supplied in work. The new estimator is parametric.[2]

As a data source I use a new, cross-sectional Danish survey of creative and innovative ability, communication ability, and reading and math ability (Hermann 2005; Klynge 2011), which provides a uniquely appropriate data set for the present purposes because it measures the extent to which individuals utilize their abilities in work in their chosen occupations. By contrast, existing data sets typically measure innate abilities using standardized test scores (e.g., Cawley, Heckman

[1] The supply of ability in work can be projected on other types of information, such as innate ability or motivation to use an ability, if such pieces of information are available.

[2] Das, Newey, and Vella (2003) consider the identification of a nonparametric selection model in which endogenous regressors are subject to selection along with the equation of interest. Their model is a reduced form model which does not build on a structural model.

and Vytlacil 1999; Heckman, Stixrud, and Urzua 2006). Moreover, the new Danish survey pinpoints the degree to which individuals use their abilities in leisure, and it is this information that my identification and estimation strategy requires. Finally, the survey can be linked to administrative data that provide precise information, for example, on wages and occupations.

The empirical results show that the individual's supply of ability should be allowed to be endogenous to correct for sample selection bias that arises from occupational choice. Essentially, for white-collar workers, the estimated return to any of the three abilities corresponds to approximately *two* years of schooling in the selection model with endogenous supply of ability. By contrast, in a conventional selection model, in which the individual's supply of ability is constant across occupations, the estimated return to communication ability corresponds to *three* years of schooling, the estimated return to reading and math ability corresponds to *two* years of schooling, and the estimated return to creative and innovative ability is close to zero and much *smaller* than the effect of one year of schooling for white-collar workers. For blue-collar workers, the estimated returns also vary across the two selection models for each ability, but the estimated returns are generally close to zero.

The paper proceeds as follows. Section II presents the theoretical version of a selection model in which the supply of ability is endogenous, while Section III presents an econometric version of the model. Section IV identifies returns to abilities within individual occupations. Section V describes the data and provides summary statistics. Section VI presents estimates of returns to abilities for both white-collar and blue-collar workers. Section VII concludes.

II. Theory

The economy has J types of occupations, with $j = 1,...,J$. For example, one occupation represents shop assistants, $j = 1$, and another represents professors, $j = 2$.[3] Each individual spends time in both work and leisure, with tasks varying across occupations and leisure activities. For example, shop assistants may be understood to "sell widgets," professors may be understood to "write research papers," and all individuals may be understood to "do housework" in their leisure time. Individuals vary in the extent to which they perform the tasks associated with given occupations or leisure. For example, some professors write many research papers, while others write few, in a given time period.

Individuals are endowed with K types of abilities, $\overline{A}_k$, with $k = 1,...,K$. One's endowment is the "stock" of his or her abilities within a given time period. For example, one type of ability is creative and innovative ability, $k = 1$, and another type of ability is communication ability, $k = 2$. The endowment of creative and innovative ability is defined as the ability to solve a previously unresolved problem on a given day. The endowment of communication ability is defined as the ability to utilize acquired knowledge in discussion on a given day. Endowment levels vary across individuals. One's ability endowment will also be referred to as "innate ability" throughout the paper.[4] Additionally, one's endowment can be understood as his or her potential or capability stock (e.g., Hartog 1977; Cunha and Heckman 2009).

[3] To build intuition, occupations and definitions of abilities are more specific in the theoretical part than in the econometric and empirical parts of this paper. For example, I refer to professors and shop assistants in the theoretical section, while I address white-collar and blue-collar workers in the later sections. White-collar workers include professors, and blue-collar workers include shop assistants. To keep the model and notation simple, I abstain from modeling task functions and distributions of innate abilities and abilities supplied, such as one finds in Heckman and Sedlacek (1985, 1990) and Gould (2002). In addition, I abstain from involving traditional explanatory variables in the income function in the theoretical part, but control for them in the empirical part of the paper.

[4] An individual's endowment level and innate ability are the individual's present stock of abilities. I do not mean to imply that this present level is necessarily the level that the person was born with.

Let A_{kj}^{w} be the supply of ability k in occupation j, and let A_{k}^{l} be the supply of ability k in leisure. The supply is the ability flow over a given time period. For example, the supply of creative and innovative ability in work is the number of previously unresolved problems solved at work on a given day, and the supply of communication ability in work is the number of known problems discussed at work on a given day. The supply of ability can also be understood as the realized value of the potential or the realized value of the capability stock.

Potential income in occupation j, Y_j, for the individual is

$$Y_j = \sum h_{kj}\left(A_{kj}^{w}\right) \tag{1}$$

where h_{kj} is the income index for ability k in occupation j. Income is increasing and strictly concave in the supply of ability, so $\partial Y_j \big/ \partial A_{kj}^{w} > 0$, $\partial^2 Y_j \big/ \partial A_{kj}^{w2} < 0$. Income is additively separable in the supply of individual abilities. Income depends on ability, but only through decisions regarding the supply of ability in work. How much the worker actually uses his abilities determines his income rather than the levels he potentially could use. For example, the number of previously unresolved research problems that a professor actually solves, not the number of research problems that he could potentially solve, determines his income.

The return to ability k in occupation j is the payoff to the ability given by $\partial Y_j \big/ \partial A_{kj}^{w}$. Following Roy (1951), abilities make workers productive; thus, returns to abilities are positive. The meaning of "return" follows common usage. Heckman, Lochner, and Todd (2006) specify the conditions under which price is in fact the rate of return in terms of schooling. Strictly speaking, the return is "the return to the supply of ability" here, rather than "the return to innate ability" as in previous studies (e.g., Roy 1951; Heckman and Sedlacek 1985, 1990; Willis 1986; Heckman and Honoré 1990; Cawley, Heckman, and Vytlacil 1999). The return to the supply of ability equals the return to innate ability only if workers use their entire ability endowment for work.

The individual faces nonpecuniary utility and disutility (i.e., "psychic cost") from using his abilities in work and in leisure. The potential individual psychic cost of working in occupation j, C_j^w, is

$$C_j^w = \sum c_{kj}^w \left(A_{kj}^w, \overline{A}_k \right) \tag{2}$$

where c_{kj}^w is the psychic cost index for ability k in occupation j. The psychic cost decreases with the supply of ability up to a reference point (i.e., $\partial C_j^w \big/ \partial A_{kj}^w < 0$ for $A_{kj}^w < \alpha_{kj}^w \overline{A}_k$, with $0 < \alpha_{kj}^w < 1$) and increases with the supply of ability beyond that point (i.e., $\partial C_j^w \big/ \partial A_{kj}^w > 0$ for $A_{kj}^w > \alpha_{kj}^w \overline{A}_k$). The function is strictly convex in the supply of ability, that is, $\partial^2 C_j^w \big/ \partial A_{kj}^{w2} > 0$. The higher the innate ability is, the higher the reference point is.

The psychic cost of leisure, C^l, for the individual is

$$C^l = \sum c_k^l \left(A_k^l, \overline{A}_k \right) \tag{3}$$

where c_k^l is the psychic cost index for ability k in leisure. The properties of the psychic cost of leisure follow those for the psychic cost of work. The reference point for the supply of an ability in leisure is $\alpha_k^l \overline{A}_k$, with $0 < \alpha_k^l < 1$.

The notion that human capital decisions involve psychic costs is consistent with the literature (e.g., Heckman, Lochner, and Todd 2006; Borghans, Duckworth, Heckman, and Weel 2008). The formal properties of the functions applied here build on the following intuition. The reference point of the psychic cost function allows an individual, engaged in a given task, to enjoy using his creative and innovative ability up to a certain point. Beyond that point, however, he begins to dislike using his creative and innovative ability in that task because it requires energy and effort. For example, a professor may enjoy solving the first seven unresolved problems in writing a research paper on a given day. Beyond that point, however, he begins to dislike solving additional unresolved research problems during that day.

The convexity assumption implies that as the professor increasingly uses his creative and innovative ability in his research on a given day, the increase in his nonpecuniary utility from using that ability decreases up to his reference point. For example, he finds it more enjoyable to solve the first research problem than the second on a given day. Beyond the reference point, his nonpecuniary disutility increases with each additional unit of ability supplied because solving each additional unresolved research problem becomes increasingly effortful on that day.

The reference point increases with one's endowment of ability. For example, with higher innate creative and innovative ability, a professor can more easily solve unresolved problems and thus will enjoy solving a larger number of unresolved problems on a given day.

The reference point can vary across occupations and leisure activities because abilities are used in different tasks and individuals enjoy using their abilities in certain tasks more than they enjoy using them in others. For example, a person may enjoy using his creative and innovative ability to solve unresolved research problems more than he enjoys using it to sell widgets or do housework.

An individual's level of innate ability is constant across occupations and leisure activities because, for example, an individual is inherently the same creative and innovative person whether he is a professor, a shop assistant, or at home on a given day. However, the amount of ability that the individual supplies may differ depending on whether he is a professor, a shop assistant, or at home on a given day. The assumption that an individual's level of innate ability is constant across occupations follows from the Roy model.

The potential utility from working in occupation j for the individual is $V_j = Y_j - C_j^w$. The utility depends on his individual income and the psychic cost of work in the occupation. I assume that the utility function is decreasing at high levels of supply of ability, such that marginal income is less than the marginal psychic cost when the supply of ability equals innate ability, that is, $h'_{kj}\left(\overline{A}_k\right) < c^{w\prime}_{kj}\left(\overline{A}_k, \overline{A}_k\right)$. This assumption helps ensure an interior solution.

The utility from leisure equals the individual's psychic cost of using his abilities in leisure activities. No income is generated from leisure activities.

The model is a partial equilibrium model, so the returns to abilities are exogenous. There is perfect competition; thus, individuals take returns as given. Each individual has perfect information and knows his potential income and supply of ability in each occupation. Individuals can move freely between occupations. The model focuses on the short run; hence, endowment levels are fixed. All individuals are employed and work the same numbers of hours.

The individual selects his type of occupation and the amount of abilities to supply in each occupation and in leisure. To select his type of occupation, he first optimizes how much of his abilities to supply in each occupation by maximizing his utility in *each* occupation. Then, he determines his utility in each occupation for the given optimal supply and maximizes his utility *across* occupations. To determine his optimal supply of abilities in leisure, he maximizes his utility in leisure. Hence, the optimal supply of ability is also called the "ability supplied" and the "ability used."

The individual selects into the occupation that yields the highest utility:

$$V^* = \max\left(V_1, ..., V_J\right) \tag{4}$$

with

$$V_j = \underset{A^w_{kj} \in \left[0, \overline{A}_k\right]}{Max} \left\{ \sum \left(h_{kj}\left(A^w_{kj}\right) - c^w_{kj}\left(A^w_{kj}, \overline{A}_k\right) \right) \right\} \tag{5}$$

The individual cannot supply a negative level of ability, and he cannot supply a higher level of ability than his innate level. These constraints on the supply of ability facilitate the development of intuition but can be omitted without the loss of an interior solution.

The optimal supply of ability in work and in leisure are interior and given by the first-order conditions arising from the assumptions I have made. The potential optimal supply of ability k in occupation j, A_{kj}^{w*}, is determined by $\partial Y_j \big/ \partial A_{kj}^{w} = \partial C_j^{w} \big/ \partial A_{kj}^{w}$. The individual supplies ability up to the point at which the marginal benefit equals the marginal cost.

The model implies that the optimal supply of ability varies across occupations when the marginal benefit or marginal cost of supply varies across occupations. For example, an individual supplies a higher level of his creative and innovative ability as a professor than as a shop assistant when the return to creative and innovative ability is higher for professors than for shop assistants and/or the psychic cost of creativity and innovation is lower when the individual performs research than when the individual sells widgets.

The optimal supply of ability k in leisure, A_k^{l*}, is determined by $\partial C^{l} \big/ \partial A_k^{l} = 0$. The individual supplies a level of ability equal to his reference point in leisure, that is, $A_k^{l*} = \alpha_k^{l} \overline{A}_k$. Although the individual receives no income from supplying his abilities in leisure, he supplies a positive level of ability because he gains nonpecuniary utility from doing so up to his reference point.

The model allows the optimal supply of ability in work to be projected on the optimal supply of ability in leisure because the optimal supply in both work and leisure depends on innate ability, $A_{kj}^{w*}\left(\overline{A}_k\right)$ and $A_k^{l*}\left(\overline{A}_k\right)$. The relationship between the optimal supply of ability in work and the optimal supply of ability in leisure is positive.[5] The higher the innate ability is, the higher the

[5] The positive relationship between the supply of ability in work and the supply of ability in leisure for each ability is supported empirically in this paper. In addition, Krueger and Schkade (2008) find positive correlations between the supply of ability in work and supply of ability in leisure with respect to social competency among American and French workers. By comparison, if a trade-off existed such that a person who employed a high level of

reference point is for the supply of ability in work and in leisure, and the higher the optimal supply of ability is in any occupation and in leisure. For example, if one individual has higher innate creative and innovative ability than another, one has a higher reference point than the other with respect to supply of creative and innovative ability, regardless of whether one is conducting research, selling widgets, or doing housework. Hence, he will supply a higher level of creative and innovative ability regardless of whether he is a professor, a shop assistant, or at home.

The choice of occupation leads to the standard selection mechanism whereby individual income is correlated with the choice of occupation. Workers in a specific occupation are not a representative sample of all workers but vary systematically from others. In the model developed here, they vary from each other in terms of innate abilities, abilities supplied, and income. Relative differences in innate abilities, psychic costs, and returns drive the choice of occupation. By comparison, in the Roy model, an individual selects into the occupation in which he obtains the highest income. Workers within a given occupation differ from a representative sample of all workers in terms of innate abilities and income. Relative differences in innate abilities and returns drive the occupational choice.

Although the model developed here builds on the Roy model, the two models are not nested. The Roy model assumes that measured abilities are constant across occupations but imposes no assumptions regarding how abilities are formed. The model developed here, by contrast, permits measured abilities to vary across occupations while imposing assumptions regarding how the abilities are formed. A general model of the two selection models would be one that permitted abilities to vary across occupations and imposed no assumptions regarding how abilities are formed. The Roy model and the selection model developed here would be special cases of that general model.

---

ability in work employed a low level of ability in leisure, the correlation between the supply of ability in work and the supply of ability in leisure would be negative.

III. Econometric Model

The econometric model parameterizes the theoretical model and accounts for heterogeneity. To be able to identify the return to ability within an occupation, I have the individual choose between two occupation types: a blue-collar occupation, $j = 1$, and a white-collar occupation, $j = 2$. In addition, income is linear in the supply of ability. The function for the psychic cost of work in an occupation includes an exclusion restriction that influences the choice of occupation but is excluded from income.

A. Structural Model

The income function in occupation $j = 1,2$, Y_j, is given by

$$Y_j = \beta_{0j} + \sum \beta_{kj} A_{kj}^w + \beta_{Xj} X + U_j^Y \tag{6}$$

where income is observed in the occupation that is chosen and unobserved in the occupation that is not chosen. The random variable A_{kj}^w represents the supply of ability k in occupation j. The supply of ability is observed in the occupation that is chosen and unobserved in the occupation that is not chosen. The random variable X is an observed control variable (or control variables). The random variable U_j^Y is unobserved and represents unobserved factors that influence income. β_{0j}, β_{kj}, and β_{Xj} are parameters. The parameter of interest is the return to ability k in occupation j, β_{kj}.

The function for the psychic cost of work in occupation j, C_j^w, is

$$C_j^w = \delta_{0j}^w + \sum \left[\left(\delta_{1kj}^w A_{kj}^w - \left(\delta_{2k}^w + \delta_{3k}^w \overline{A}_k + v_{kj}^w \right) \right)^2 \right] + \delta_{Zj}^w Z + U_j^{C^w} \tag{7}$$

where the psychic cost is unobserved. The random variable $\overline{A}_k$ represents innate ability of type k, which is unobserved.

The random variable Z is an observed exclusion restriction. Following the previous literature, I impose that the occupation type of the individual's father influences the individual's choice of occupation but does not affect the individual's income, and I apply that piece of information as the exclusion restriction (Cawley, Heckman, and Vytlacil 1999; Nielsen, Simonsen, and Verner 2004). Thus, Z indicates whether the father was a white-collar worker. The assumption underlying this restriction is that the individual has an aversion to blue-collar work if his father was a white-collar worker. This aversion increases his psychic cost of performing blue-collar work, that is, $\delta_{Z1}^{w} > 0$, and influences his choice of occupation but not his income. Vice versa, the individual's psychic cost of performing white-collar work decreases if his father was a white-collar worker, that is, $\delta_{Z2}^{w} < 0$.

The random variable ν_{kj}^{w} is unobserved. This component allows the psychic cost of supplying ability k in occupation j to vary across workers for a given supply of ability and given innate ability, for example, owing to unobserved working conditions. The random variable $U_j^{C^w}$ is unobserved and represents unobserved factors that influence the psychic cost of working in occupation j. $\delta_{0j}^{w}, \delta_{1kj}^{w}, \delta_{2k}^{w}, \delta_{3k}^{w}$, and δ_{Zj}^{w} are parameters.

The psychic cost of supplying ability k in occupation j decreases as the supply of the ability increases up to the reference point, $\left(\delta_{2k}^{w} + \delta_{3k}^{w}\overline{A}_k + \nu_{kj}^{w}\right)\big/\delta_{1kj}^{w}$, and increases beyond that point. Assuming $\delta_{1kj}^{w} > 0$, the psychic cost is strictly convex in the supply of ability. Assuming $\delta_{3k}^{w} > 0$, and thus $\delta_{3k}^{w}\big/\delta_{1kj}^{w} > 0$, an increase in innate ability increases the reference point.

As for the unobserved random variables in the income function and in the function for the psychic cost of work in the two occupations, $U_j^Y, U_j^{C^w}$, and ν_{kj}^w, I assume they have a mean zero multivariate normal distribution with

$$\left(\left\{U_j^Y, U_j^{C^w}, \nu_{kj}^w\right\}_{j=1,2;k=1,\dots,K}\right) \sim N\left(0, \Sigma_{U_j^Y, U_j^{C^w}, \nu_{kj}^w}\right)$$

where $\Sigma_{U_j^Y, U_j^{C^w}, \nu_{kj}^w}$ is the variance-covariance matrix.

The function for the psychic cost in leisure, C^l, is

$$C^l = \delta_0^l + \sum\left[\left(\delta_{1k}^l A_k^l - \left(\delta_{2k}^l + \delta_{3k}^l \overline{A}_k\right)\right)^2\right] + U^{C^l} \quad (8)$$

where the psychic cost is unobserved. The random variable A_k^l represents the supply of ability k in leisure, which is observed. The random variable U^{C^l} is unobserved and captures unobserved factors that influence the psychic cost in leisure. $\delta_0^l, \delta_{1k}^l, \delta_{2k}^l$, and δ_{3k}^l are parameters.

The reference point in leisure is given by $\left(\delta_{2k}^l + \delta_{3k}^l \overline{A}_k\right) / \delta_{1k}^l$. Assuming $\delta_{1k}^l > 0$, the psychic cost is strictly convex in the supply of ability. Assuming $\delta_{3k}^l > 0$, and thus $\delta_{3k}^l / \delta_{1k}^l > 0$, the higher one's innate ability is, the higher one's reference point is.

The utility function from working in occupation j, V_j, is

$$V_j = Y_j - C_j^w \quad (9)$$

where utility is unobserved. Assuming $\beta_{kj} < 2 \cdot \delta_{1kj}^w \left[\left(\delta_{1kj}^w - \delta_{3k}^w\right)\overline{A}_k - \delta_{2k}^w - \nu_{kj}^w\right]$, the utility from working in an occupation decreases at high levels of the supply of ability.

Let D be an indicator that equals one if the individual decides to become a white-collar worker and zero if he decides to become a blue-collar worker:

$$D = 1 \text{ if } S \geq 0; \quad D = 0 \text{ otherwise} \quad (10)$$

The random variable for the choice of occupation, D, is observed. The worker selects the occupation that yields the highest utility. The selection equation, S, is then

$$S = V_2 - V_1 \tag{11}$$

where the random variable S is unobserved. This variable measures the net gain in utility from choosing white-collar work over blue-collar work.

B. Optimal Supply of Ability

The optimal supply of ability k in occupation j, A_{kj}^{w*}, is determined by $\partial Y_j / \partial A_{kj}^w = \partial C_j^w / \partial A_{kj}^w$ and equals the following expression, given (6) and (7):

$$A_{kj}^{w*} = \frac{\beta_{kj}}{2 \cdot \delta_{1kj}^w} + \frac{\delta_{2k}^w}{\delta_{1kj}^w} + \frac{\delta_{3k}^w}{\delta_{1kj}^w} \overline{A}_k + \frac{1}{\delta_{1kj}^w} v_{kj}^w \tag{12}$$

The optimal supply of ability varies across occupations if either the return, β_{kj}, or the psychic cost, δ_{1kj}^w, associated with use of the ability vary across occupations. Additionally, the optimal supply varies across occupations if the unobserved random variable v_{kj}^w from the psychic cost of supplying ability k varies across occupations. The ability supplied in work is positively correlated with innate ability owing to the previous assumptions of $\delta_{3k}^w > 0$ and $\delta_{1kj}^w > 0$.

The optimal supply of ability k in leisure, A_k^{l*}, is determined by $\partial C^l / \partial A_k^l = 0$. Given (8), it is:

$$A_k^{l*} = \frac{\delta_{2k}^l}{\delta_{1k}^l} + \frac{\delta_{3k}^l}{\delta_{1k}^l} \overline{A}_k \tag{13}$$

The ability supplied in leisure is positively correlated with innate ability owing to the previous assumptions of $\delta_{3k}^l > 0$ and $\delta_{1k}^l > 0$.

The optimal supply of ability in work can now be projected on the optimal supply of ability in leisure through innate ability. Solving (13) for innate ability

$$\overline{A}_k = -\frac{\delta^l_{2k}}{\delta^l_{3k}} + \frac{\delta^l_{1k}}{\delta^l_{3k}} A^{l*}_k \tag{14}$$

and substituting (14) into (12) yields the projection of the optimal supply of ability in work on the optimal supply of ability in leisure:

$$A^{w*}_{kj} = \theta_{0kj} + \theta_{1kj} A^{l*}_k + \varepsilon_{kj} \tag{15}$$

where the random variable ε_{kj} is given by

$$\varepsilon_{kj} = \frac{1}{\delta^w_{1kj}} v^w_{kj} \tag{16}$$

This variable is unobserved and represents the unobserved heterogeneity that influences the optimal supply of ability in each occupation. θ_{0kj} and θ_{1kj} are parameters defined as

$$\theta_{0kj} = \frac{\beta_{kj}}{2 \cdot \delta^w_{1kj}} + \frac{\delta^w_{2k}}{\delta^w_{1kj}} - \frac{\delta^w_{3k} \cdot \delta^l_{2k}}{\delta^w_{1kj} \cdot \delta^l_{3k}} \tag{17}$$

$$\theta_{1kj} = \frac{\delta^w_{3k} \cdot \delta^l_{1k}}{\delta^w_{1kj} \cdot \delta^l_{3k}} \tag{18}$$

The parameter θ_{1kj} is positive owing to the previous assumptions, namely, $\delta^w_{1kj} > 0$, $\delta^w_{3k} > 0$, $\delta^l_{1k} > 0$, and $\delta^l_{3k} > 0$. Thus, the relationship between the optimal supply of ability in each occupation and the optimal supply of ability in leisure is positive for each type of ability.

C. Income, Psychic Cost, and Selection Equation at the Optimum

Income in occupation j at the optimum is

$$Y_j = \beta_{0j} + \sum \beta_{kj} A^{w*}_{kj} + \beta_{Xj} X + U^Y_j \tag{19}$$

The psychic cost of work in occupation j at the optimum is

$$C_j^w = \delta_{0j}^w + \sum\left[\left(\delta_{1kj}^w A_{kj}^{w*} - \left(\delta_{2k}^w + \delta_{3k}^w \overline{A}_k + \nu_{kj}^w\right)\right)^2\right] + \delta_{Zj}^w Z + U_j^{C^w} \tag{20}$$

The psychic cost of leisure at the optimum is

$$C^l = \delta_0^l + \sum\left[\left(\delta_{1k}^l A_k^{l*} - \left(\delta_{2k}^l + \delta_{3k}^l \overline{A}_k\right)\right)^2\right] + U^{C^l} \tag{21}$$

Substituting (9), (12), (14), (19), and (20) into (11) yields the selection equation at the optimum:

$$S = \mu_0 + \sum \mu_k A_k^{l*} + \mu_X X + \mu_Z Z - V \tag{22}$$

where the random variable V is defined by

$$V = \left(U_1^Y - U_1^{C^w}\right) - \left(U_2^Y - U_2^{C^w}\right) + \sum\left(\frac{\beta_{k1}}{\delta_{1k1}^w}\nu_{k1}^w - \frac{\beta_{k2}}{\delta_{1k2}^w}\nu_{k2}^w\right) \tag{23}$$

This variable is unobserved and represents the unobserved heterogeneity that influences the choice of occupation. μ_0, μ_k, μ_X, and μ_Z are parameters defined by

$$\mu_0 = \beta_{02} - \beta_{01} + \delta_{01}^w - \delta_{02}^w + \sum\left[\frac{1}{4}\left(\beta_{k1}^2 - \beta_{k2}^2\right) + \frac{1}{2}\left(\frac{\beta_{k2}^2}{\delta_{1k2}^w} - \frac{\beta_{k1}^2}{\delta_{1k1}^w}\right) + \left(\delta_{2k}^w - \frac{\delta_{3k}^w \cdot \delta_{2k}^l}{\delta_{3k}^l}\right)\left(\frac{\beta_{k2}}{\delta_{1k2}^w} - \frac{\beta_{k1}}{\delta_{1k1}^w}\right)\right] \tag{24}$$

$$\mu_k = \frac{\delta_{3k}^w \cdot \delta_{1k}^l}{\delta_{3k}^l}\left(\frac{\beta_{k2}}{\delta_{1k2}^w} - \frac{\beta_{k1}}{\delta_{1k1}^w}\right) \tag{25}$$

$$\mu_X = \beta_{X2} - \beta_{X1} \tag{26}$$

$$\mu_Z = \delta_{Z1}^w - \delta_{Z2}^w \tag{27}$$

The parameter μ_Z is positive owing to the previous assumptions, namely, $\delta_{Z1}^w > 0$ and $\delta_{Z2}^w < 0$.

IV. Identification of Return to Ability within an Occupation

A. Reduced Form Model and Assumptions

I identify the return to ability k in occupation j, β_{kj}, using the selection model in reduced form:

$$D = 1 \text{ if } S \geq 0; \quad D = 0 \text{ otherwise} \tag{28}$$

$$S = \mu_0 + \sum \mu_k A_k^{l*} + \mu_X X + \mu_Z Z - V \tag{29}$$

$$Y_j = \beta_{0j} + \sum \beta_{kj} A_{kj}^{w*} + \beta_{Xj} X + U_j^Y \tag{30}$$

$$A_{kj}^{w*} = \theta_{0kj} + \theta_{1kj} A_k^{l*} + \varepsilon_{kj} \tag{31}$$

These equations are, respectively, the selection equation, the income equation, and the equation for the supply of ability in work in reduced form.

Let measured income, Y, and the measured supply of ability k in work, A_k^{w*}, be defined as

$$Y = DY_2 + (1 - D) Y_1 \tag{32}$$

$$A_k^{w*} = DA_{k2}^{w*} + (1 - D) A_{k1}^{w*} \tag{33}$$

The observed random variables are $D, Y, A_k^{w*}, A_k^{l*}, X$, and Z, and the unobserved random variables are V, U_j^Y, and ε_{kj}.

I make the following assumptions for the identification:[6]

A.1: The observed random variables A_k^{l*}, X, and Z, are statistically independent of the unobserved random variables V, U_j^Y, and ε_{kj}, respectively.

[6] Assumptions A.2 and A.3 follow from the structural model in which I assume that the unobserved random variables in the income function and in the function for the psychic cost of work in the two occupations have a mean zero multivariate normal distribution.

A.2: The unobserved random variables in the selection equation and in the income equation have a bivariate normal distribution with

$$\begin{pmatrix} V \\ U_j^Y \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_V^2 & \sigma_{V,U_j^Y} \\ \sigma_{V,U_j^Y} & \sigma_{U_j^Y}^2 \end{pmatrix} \right)$$

where σ_{V,U_j^Y} is the covariance and σ_V^2 and $\sigma_{U_j^Y}^2$ are the variances.

A.3: The unobserved random variables in the selection equation and in the equation for the supply of ability in work have a bivariate normal distribution with

$$\begin{pmatrix} V \\ \varepsilon_{kj} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_V^2 & \sigma_{V,\varepsilon_{kj}} \\ \sigma_{V,\varepsilon_{kj}} & \sigma_{\varepsilon_{kj}}^2 \end{pmatrix} \right)$$

where $\sigma_{V,\varepsilon_{kj}}$ is the covariance and $\sigma_{\varepsilon_{kj}}^2$ is the variance.

B. Identification for White-Collar Workers

A three-step procedure leads to identification of the return to ability k for white-collar workers, β_{k2}. The first step involves the conditional expectation of the choice of occupation (or propensity score) for the full sample. The second step addresses the conditional expectation of the supply of ability in work for white-collar workers. The third step uses the conditional expectation of income for white-collar workers.

Formally, the three conditional expectations are as follows (see Appendix A for details). The conditional expectation of the choice of occupation, D, given the observed variables, $\varpi = \left\{ A_k^{l*}, X, Z \right\}$, is

$$\mathrm{E}\left[D \mid \varpi \right] = \mathrm{P}\left(D = 1 \mid \varpi \right) = \Phi\left(g\left(\varpi \right) / \sigma_V \right) \tag{34}$$

where $\Phi\left(g\left(\varpi\right)/\sigma_V\right)$ is the standard normal cumulative distribution function evaluated at point $g\left(\varpi\right)/\sigma_V$ and

$$g\left(\varpi\right)=\mu_0+\sum\mu_k A_k^{l*}+\mu_X X+\mu_Z Z \tag{35}$$

I derive (34) from (28) and (29) and assumptions A.1 and A.2.

The conditional expectation of the supply of ability k in work for white-collar workers, A_{k2}^{w*}, given the observed variables ϖ and the choice of being a white-collar, $D=1$, is

$$E\left[A_{k2}^{w*}\mid\varpi,D=1\right]=\theta_{0k2}+\theta_{1k2}A_k^{l*}+\rho_{V,\varepsilon_{k2}}\sigma_{\varepsilon_{k2}}\lambda_2 \tag{36}$$

where λ_2 is the inverse Mills ratio for white-collar workers determined by

$$\lambda_2\equiv\lambda_2\left(g\left(\varpi\right)/\sigma_V\right)=-\varphi\left(g\left(\varpi\right)/\sigma_V\right)/\Phi\left(g\left(\varpi\right)/\sigma_V\right)$$

and $\varphi\left(g\left(\varpi\right)/\sigma_V\right)$ is the standard normal probability density function evaluated at point $g\left(\varpi\right)/\sigma_V$. The parameter $\rho_{V,\varepsilon_{k2}}$ is the correlation, and $\sigma_{\varepsilon_{k2}}$ and σ_V are the standard deviations. The conditional expectation corrects for the selection bias that arises if unobserved factors influence both the choice of occupation and the supply of ability in work for white-collar workers. I derive (36) from (28), (29), and (31) and assumptions A.1 and A.3.

The conditional expectation of income for white-collar workers, Y_2, given the observed variables ϖ and the choice of being a white-collar worker, $D=1$, is

$$E\left[Y_2\mid\varpi,D=1\right]=\gamma_{02}+\sum\gamma_{k2}A_k^{l*}+\beta_{X2}X+\left[\sum\beta_{k2}\rho_{V,\varepsilon_{k2}}\sigma_{\varepsilon_{k2}}+\rho_{V,U_2^Y}\sigma_{U_2^Y}\right]\lambda_2 \tag{37}$$

where γ_{20} and γ_{2k} are parameters defined as

$$\gamma_{02}=\beta_{02}+\sum\beta_{k2}\theta_{0k2} \tag{38}$$

$$\gamma_{k2}=\beta_{k2}\theta_{1k2} \tag{39}$$

The conditional expectation corrects for the selection bias that arises if unobserved factors influence both the choice of occupation and the income of white-collar workers. The selection effect can be decomposed into two effects. The first is an ability supply selection effect, $\sum \beta_{k2}\rho_{V,\varepsilon_{k2}}\sigma_{\varepsilon_{k2}}$, which captures the influence on income through the supply of ability in work. The second is a conventional selection effect, $\rho_{V,U_2^Y}\sigma_{U_2^Y}$, which captures the direct effect on income of unobserved factors. I derive (37) from (28) to (31) and assumptions A.1 to A.3.

The three conditional expectations form the basis for the following identification procedure.

Three-Step Identification Procedure for White-Collar Workers

Step 1: Estimate a probit specification of the selection equation for the full sample:

$$P(D=1)=\Phi\left(\frac{\mu_0}{\sigma_V}+\sum\frac{\mu_k}{\sigma_V}A_k^{l*}+\frac{\mu_X}{\sigma_V}X+\frac{\mu_Z}{\sigma_V}Z\right) \tag{40}$$

The first step yields values for $\Phi\left(g(\varpi)/\sigma_V\right)$ and $\varphi\left(g(\varpi)/\sigma_V\right)$ and hence λ_2.

Step 2: Estimate the equation for ability supply in work for white-collar workers:

$$A_{2k}^{w*}=\theta_{0k2}+\theta_{1k2}A_k^{l*}+\rho_{V,\varepsilon_{k2}}\sigma_{\varepsilon_{k2}}\lambda_2+\upsilon_2 \tag{41}$$

where the unobserved random variable υ_2 is defined by $\upsilon_2=A_{k2}^{w*}-E\left[A_{k2}^{w*}\mid\varpi,D=1\right]$. The second step yields values for $\theta_{0k2},\theta_{1k2}$, and $\left(\rho_{V,\varepsilon_{k2}}\sigma_{\varepsilon_{k2}}\right)$.

Step 3: Estimate the income equation for white-collar workers:

$$Y_2=\gamma_{02}+\sum\gamma_{k2}A_k^{l*}+\beta_{X2}X+\left(\sum\beta_{k2}\rho_{V,\varepsilon_{k2}}\sigma_{\varepsilon_{k2}}+\rho_{V,U_2^Y}\sigma_{U_2^Y}\right)\lambda_2+\kappa_2 \tag{42}$$

where the unobserved random variable κ_2 is defined by $\kappa_2=Y_2-E\left[Y_2\mid\varpi,D=1\right]$. The third step yields values for γ_{02}, γ_{k2}, β_{X2}, and $\left(\sum\beta_{k2}\rho_{V,\varepsilon_{k2}}\sigma_{\varepsilon_{k2}}+\rho_{V,U_2^Y}\sigma_{U_2^Y}\right)$.

The parameter of interest is identified from the parameters identified above. That is, the return to ability k for white-collar workers, β_{k2}, is determined from the parameter γ_{k2}, identified in step 3, and the parameter θ_{1k2}, identified in step 2, as $\gamma_{k2} = \beta_{k2}\theta_{1k2}$ (see (39)).

The conventional selection effect, $\rho_{V,U_2^Y}\sigma_{U_2^Y}$, is determined from the parameter $\sum \beta_{k2}\rho_{V,\varepsilon_{k2}}\sigma_{\varepsilon_{k2}} + \rho_{V,U_2^Y}\sigma_{U_2^Y}$, identified in step 3, the parameter $\rho_{V,\varepsilon_{k2}}\sigma_{\varepsilon_{k2}}$, determined in step 2, and the parameter β_{k2}, identified above.

Testing whether the coefficient on λ_2 in step 3 equals zero, $\left(\sum \beta_{k2}\rho_{V,\varepsilon_{k2}}\sigma_{\varepsilon_{k2}} + \rho_{V,U_2^Y}\sigma_{U_2^Y}\right) = 0$, shows whether selection bias is present in the returns to abilities for white-collar workers, β_{k2}.

C. Identification for Blue-Collar Workers

The return to ability k for blue-collar workers, β_{k1}, is identified through a three-step procedure similar to that employed for white-collar workers; however, with the replacement

$$\lambda_1 \equiv \lambda_1\left(g(\varpi)/\sigma_V\right) = \varphi\left(g(\varpi)/\sigma_V\right)\Big/\left(1-\Phi\left(g(\varpi)/\sigma_V\right)\right)$$

V. Data and Summary Statistics

A. Data

I use a cross-sectional Danish survey of individual abilities linked to administrative data.

Description of the Survey

The Danish survey, The National Competence Account, measures the extent to which individuals use 10 abilities in work and in leisure. The measure of abilities are: creative and innovative ability, communication ability, reading and math ability, self-management ability, social competency, learning ability, civic competency, health awareness, environmental awareness, and intercultural awareness (Hermann 2005; Klynge 2011).

The survey covers abilities considered to be key competencies for a successful life and a well-functioning society in a knowledge-based economy. The 10 abilities are selected and defined theoretically by the Organisation for Economic Co-operation and Development (OECD) in the project Definition and Selection of Key Competencies: Theoretical and Conceptual Foundations, which was undertaken by the OECD in 1997–2003 (Salganik and Rychen 2003).

The National Competence Account project, conducted from 2001 to 2005, developed the survey and measured abilities empirically in Denmark. The project conducted a pilot study that led to the development of the main study. I use the data from the main study in this paper.

Statistics Denmark collected the data from the main study in winter 2003–2004 from a representative sample of individuals aged 20 to 64 who were living in Denmark on November 1st, 2003. The survey was conducted through telephone interviews in two rounds. Of an initial sample size of 7,953, a total of 5,170 individuals responded to the survey.

I employ abilities that meet two criteria. First, as the selection model focuses on productive abilities and thus abilities with positive returns, the abilities considered must have positive returns for white-collar and blue-collar workers in a basic wage regression. Second, I must be able to create an index of the supply of ability in leisure, as my identification and estimation strategy requires such information. Creative and innovative ability, communication ability, and reading and math ability meet these criteria (Klynge 2011).

The three abilities are defined as follows (see Appendix B for details). The creative and innovative ability is defined as the ability to create new products or services. It is concerned with the ability to create new applications of existing technology, new concepts, new problem solutions, and new knowledge that differ from what already exists.

The communication ability is defined as the ability to argue one's own opinion and to understand others. This involves the ability to manage appropriate communication methods and tools for sharing information with other people.

The reading and math ability pertains to reading and math comprehension. Reading comprehension comprises the ability to understand, interpret, and reflect upon written materials. Math comprehension concerns the capacity to identify, understand, and engage in mathematics.

Administrative Data

The administrative data provide information on each individual's wage, occupation type, education, labor market experience, gender, and father's occupation type. Wage measures the hourly average wage. Education and labor market experience are measured in years. Data for all variables other than father's occupation type are from the winter 2003–2004. Data on fathers' occupation type is from 1980, the first year for which Danish administrative data are available. A personal identifier code enables the administrative data to be linked to the survey data for each individual.

Sample

The sample includes 1,888 individuals who are employees, and for whom the administrative data record positive and reliable hourly wages and information on fathers' occupations.

B. Summary Statistics

Figures 1 and 2 present histograms for abilities supplied in work and in leisure by the 1,159 white-collar workers and the 729 blue-collar workers in the sample. White-collar workers include managers, professionals, clerks, technicians, and associated professionals. Blue-collar workers include service workers, shop and market sales workers, skilled agricultural and fishery workers, craft and related trades workers, plant and machine operators and assemblers, and elementary occupations. The unit of measurement for supply of a given ability is one standard deviation (see Appendix C for details).

The two figures show that each individual supplies the three abilities in both work and leisure. For example, every white-collar and every blue-collar worker uses their creative and innovative ability in work. The exceptions are creative and innovative ability in leisure and reading and math ability in work, for which some individuals supply the minimum levels possible. The two figures also illustrate that the supply of ability in both work and leisure varies across individuals. For example, some white-collar workers supply a higher level of creative and innovative ability in work than other white-collar workers. The same holds for blue-collar workers.

Table 1 lists the means and correlations for abilities supplied in work and in leisure for both white-collar and blue-collar workers. On average, white-collar workers supply every type of ability to a greater extent in both work and leisure than do blue-collar workers. The correlations between the supply of ability in work and the supply of ability in leisure are positive and significant for each ability within an occupation. For example, white-collar workers who are likely to supply a high level of creative and innovative ability in work tend to supply a high level of the creative and innovative ability in leisure. The same holds for blue-collar workers. The correlations are higher for creative and innovative ability and communication ability than for reading and math ability in every occupation.

Table 2 presents the means for the outcome and background variables. The results show that white-collar workers differ from blue-collar workers in other dimensions, in addition to the levels of abilities supplied. On average, white-collar workers have a higher hourly wage (204 DKK (36 USD)) than blue-collar workers (166 DKK (30 USD)), as well as a higher level of education (13.6 years versus 10.6 years). By contrast, on average, the level of work experience is roughly the same between white-collar workers and blue-collar workers (close to 14.5 years). The share of males is lower among white-collar workers (42 percent) than among blue-collar workers (65 percent). Finally, white-collar workers are more likely (70 percent) than blue-collar workers (50 percent) to have fathers who were white-collar workers.

VI. Estimation Results

I estimate the returns to creative and innovative ability, communication ability, and reading and math ability for white-collar and blue-collar workers based on the selection model with endogenous supply of ability. I compare the results with the estimated returns from a basic wage regression and a conventional selection model. A basic wage regression model ignores sample selection. A conventional selection model corrects for sample selection but assumes that supply of ability is constant across occupations. A selection model in which supply of ability is endogenous corrects for sample selection and allows for supply of ability to vary across occupations.

In the tables the wage is the log of the hourly wage. Occupation type is a dummy variable indicating whether an individual is a white-collar worker. Father's occupation is also a dummy variable indicating whether an individual's father was a white-collar worker. Control variables include education, labor market experience, and gender, which are standard explanatory variables in a wage regression based on the Mincer (1974) model. The unit for education and labor market experience is one year. Experience squared is normed by 100. The unit for ability supplied is one standard deviation (see Appendix C for details).

I first report results of the initial steps used to obtain the estimated returns to abilities in the two selection models. I then present and compare the returns to abilities for all three models. For brevity, I focus on discussing the estimates of the abilities, the exclusion restriction, and the selection corrections.

A. Results from the Three Steps in the Selection Model with Endogenous Supply of Ability

The three-step procedure yields the returns to abilities for white-collar and blue-collar workers in the selection model with endogenous supply of ability. Note that the explanatory variables in the three steps are the abilities supplied in leisure, rather than the abilities supplied in work, as the abilities supplied in leisure substitute for the partially observed abilities supplied in work.

Table 3 shows the results from the first step in which a probit model of the selection equation for the full sample is estimated. Father's occupation type has a positive and statistically significant effect, with a coefficient of 0.283. An individual is more likely to be white-collar if his or her father was a white-collar worker. The three abilities (as measured in leisure) have positive and statistically significant effects on selection, with approximately the same levels for each ability, close to 0.187. The greater an individual's supplied level of any of the three abilities is, the more likely he or she is to be a white-collar worker.

Table 4 reports the results from the second step, in which the conditional expectation of each ability supplied in work in each occupation is estimated. For white-collar workers, a one-standard-deviation increase in creative and innovative ability supplied in leisure is associated with a 0.725-standard-deviation increase in quantity supplied in work. For blue-collar workers, the coefficient for creative and innovative ability is somewhat lower, at 0.629. The estimated coefficient for communication ability is 0.405 for white-collar workers and 0.243 for blue-collar workers. The estimated coefficient for reading and math ability is 0.081 for the two occupation types. The estimates are statistically significant in all cases but one. The exception is the estimate for reading and math ability, which is statistically insignificant for blue-collar workers. This lack of significance may be due to the smaller sample size, as the level of the estimate is equal to the one for white-collar workers.

The positive estimates for abilities supplied in leisure for each occupation support the model prediction of a positive relationship between ability supplied in work and ability supplied in leisure, that is, $\theta_{1kj} > 0$ (see (15)).

The estimates of abilities supplied in leisure vary across occupations for creative and innovative ability and communication ability. This finding is consistent with the model that allows the reference point for the supply of an ability to vary across occupations. To see this variation, note that the parameter for ability supplied in leisure is defined by $\theta_{1kj} = \delta_{3k}^{w} \cdot \delta_{1k}^{l} \big/ \delta_{1kj}^{w} \cdot \delta_{3k}^{l}$ (see (18)). The variation in θ_{1kj} across occupations, $j = 1, 2$, implies that the parameter δ_{1kj}^{w} varies across occupations. As a result, the reference point for the supply of an ability in work, $\left(\delta_{2k}^{w} + \delta_{3k}^{w}\overline{A}_{k} + v_{kj}^{w}\right) \big/ \delta_{1kj}^{w}$, can vary across occupations.

Reading and math ability has the lowest positive estimates with respect to θ_{1kj} in the two occupation types when compared with creative and innovative ability. Reading and math ability also has the lowest positive correlation in the summary statistics.

The estimates of λ_{j} are statistically significant in five of six cases and insignificant only in the case of reading and math ability for white-collar workers. Generally, correcting for the sample selection bias in the second step is important. Unobserved factors influence both the choice of occupation and the supply of abilities in work in each occupation.

Table 5 presents the results of the third step, in which the conditional expectation of income for each occupation type is estimated. I abstain from discussing the estimates of the abilities in this step, as they cannot be interpreted per se (see (39) and (42)). I include the estimates of the selection corrections and control variables in Tables 6 and 7 below (see Subsections C and D).

I obtain the estimated returns to abilities in each occupation from the estimates for the abilities from the second and third step. Tables 6 and 7 present the results regarding the estimated returns to abilities, and I will postpone the discussion on these results until then.

B. Results from Two-Step Procedure Applied to the Conventional Selection Model

In a conventional selection model the returns to abilities in each occupation are obtained through a two-step procedure. Note that the explanatory variables are the abilities supplied in work, rather than in leisure, as the model assumes supply of abilities to be constant across occupations.

Table 3 shows the results for the first step, in which a probit specification of the selection equation is estimated for the full sample. The estimated coefficient for father's occupation type, 0.332, is positive and statistically significant. Individuals tend to be white-collar workers if their father was a white-collar worker. Communication ability and reading and math ability also have positive coefficients, 0.897 and 0.537, respectively. The greater an individual's supplied level of either of these two abilities is, the more likely he or she is to be a white-collar worker. The coefficient for creative and innovative ability, by contrast, is negative, -0.135. A person with a higher supplied level of creative and innovative ability is more likely to be a blue-collar worker. The estimates for the three abilities are statistically significant in all cases.

The results in Table 3 indicate that whether the supply of ability is allowed to be endogenous in the first step affects the results of the models. The estimates vary quantitatively across the two selection models. Moreover, the results vary qualitatively for creative and innovative ability, which has a positive coefficient in the selection model with endogenous supply of ability and a negative coefficient in the conventional selection model. The latter negative estimate would seem counterintuitive, given that white-collar workers, on average, supply higher levels of creative and innovative ability in work than do blue-collar workers, based on the summary statistics.

In the second step, the conditional expectation of income in each occupation, taking sample selection into account, is estimated. Tables 6 and 7 below include the results of the second step, and I will postpone the discussion on these results until then.

C. Returns for White-Collar Workers

Table 6 presents the estimated returns to abilities for white-collar workers, based on a basic wage regression, a conventional selection model, and a selection model with endogenous supply of ability.

Basic Wage Regression

In the basic wage regression model, the estimate for creative and innovative ability is small, 0.7 percent, and statistically insignificant. By contrast, the estimated return to communication ability is large, 12.1 percent, and statistically significant. Thus, a white-collar worker who employs his communication ability at a level one standard deviation above the mean obtains a 12.1 percent higher wage than a white-collar worker who employs his communication ability at the mean level. The estimated return to reading and math ability is also large, 10.2 percent, and statistically significant.

Conventional Selection Model

In the conventional selection model, the estimate for creative and innovative ability is statistically insignificant and is even closer to zero, 0.2 percent, than that in the basic wage regression. Communication ability has a slightly higher estimated return, 12.4 percent, while reading and math ability has a lower estimated return, 8.6 percent, in the conventional selection model than in the basic wage regression model. As in the basic wage regression model, the estimated returns to the two abilities are statistically significant.

The results differ quantitatively from those of the basic wage regression owing to sample selection bias. The test for sample selection bias, a nonzero estimate of λ_2, rejects the hypothesis that the parameter equals zero. The significant estimate implies that unobserved factors influence both the choice of occupation and the wage of white-collar workers. Additionally, the result indicates that taking the model one step further by allowing endogenous supply of ability may be important to avoid sample selection bias.

Selection Model with Endogenous Supply of Ability

The results of the selection model with endogenous supply of ability stand in sharp contrast to those of the two previous models. In this specification, the estimated return to creative and innovative ability is reasonably large, 5.6 percent, and statistically significant. Thus, a white-collar worker who employs his creative and innovative ability at a level one standard deviation above the mean obtains a 5.6 percent higher wage than a white-collar worker who employs his creative and innovative ability at the mean level. The estimated returns to communication ability and reading and math ability are also reasonably large, 5.6 percent and 7.7 percent, respectively, but not as large as those in the conventional selection model. The returns to the two abilities are estimated imprecisely and thus are statistically insignificant.

As in the conventional selection model, the results from the selection model with endogenous supply of ability show that correcting for sample selection bias due to occupational choice is important. The test for sample selection bias, a nonzero estimate of λ_2, rejects the hypothesis that the parameter equals zero. Unobserved factors influence both the choice of occupation and the wage of white-collar workers.

A decomposition of the selection effect in the selection model with endogenous supply of ability, $\sum \beta_{k2}\rho_{V,\varepsilon_{k2}}\sigma_{\varepsilon_{k2}} + \rho_{V,U_2^Y}\sigma_{U_2^Y}$ (see (42)), shows that the significance arises from the total ability supply selection effect, $\sum \beta_{k2}\rho_{V,\varepsilon_{k2}}\sigma_{\varepsilon_{k2}}$, while the conventional selection effect, $\rho_{V,U_2^Y}\sigma_{U_2^Y}$, is insignificant. The unobserved factors that influence both the choice of occupation and the wage of white-collar workers influence wages through supply of abilities in white-collar work but have no direct influence on wages.

Further decomposition of the selection effect into the effects on each individual ability, $\beta_{k2}\rho_{V,\varepsilon_{k2}}\sigma_{\varepsilon_{k2}}$, shows that the selection effect influences creative and innovative ability and communication ability, which have significant effects, but not reading and math ability, which has an insignificant effect. These results follow from the second step, which shows that only the estimates for creative and innovative ability and communication ability significantly suffer from selection bias.

D. Returns for Blue-Collar Workers

Table 7 reports the estimated returns to abilities for blue-collar workers from a basic wage regression, a conventional selection model, and a selection model with endogenous supply of ability.

The table shows that the estimated returns vary across the three models for each ability. For example, the estimated return to creative and innovative ability is 1 percent in the basic wage regression, 0.8 percent in the conventional selection model, and 0.1 percent in the selection model with endogenous supply of ability. However, the estimated returns are generally close to zero and statistically insignificant. One exception is reading and math ability, which is estimated to have a negative return in the selection model with endogenous supply of ability. Nevertheless, the return to reading and math ability is imprecisely estimated in this model and cannot be rejected as being zero.

Even if the returns to the three abilities are zero, the blue-collar workers still supply each of the three abilities in work, based on the summary statistics. These findings are consistent with the model prediction that every individual supplies ability to some extent because they gain nonpecuniary utility from doing so.

The two selection models both show that the estimates from the basic wage regression should be corrected for sample selection bias, with the estimates of λ_1 being statistically significant in both selection models.

A decomposition of the selection effect in the selection model with endogenous supply of ability, $\sum \beta_{k1}\rho_{V,\varepsilon_{k1}}\sigma_{\varepsilon_{k1}} + \rho_{V,U_1^Y}\sigma_{U_1^Y}$, shows that the significance arises from the conventional selection effect, $\rho_{V,U_1^Y}\sigma_{U_1^Y}$. The unobserved factors that influence the choice of occupation and the wage of blue-collar workers have a direct influence on wages.

The part of the selection effect related to the total supply of ability, $\sum \beta_{k1}\rho_{V,\varepsilon_{k1}}\sigma_{\varepsilon_{k1}}$, and to each individual ability, $\beta_{k1}\rho_{V,\varepsilon_{k1}}\sigma_{\varepsilon_{k1}}$, is statistically insignificant. The unobserved factors that influence the choice of occupation and the wage of blue-collar workers have no influence on wages through supply of abilities in blue-collar work. The insignificant estimates are due to the generally small and statistically insignificant estimated returns to the three abilities, β_{k1}, as the second step shows the presence of a selection bias for each ability, $\rho_{V,\varepsilon_{k1}}\sigma_{\varepsilon_{k1}} \neq 0$.

VII. Conclusion

I have developed a selection model in which the individual's supply of ability is endogenous and subject to selection along with occupation. The model is motivated by the potential variation in the individual's decision regarding the amount of ability to supply across occupations. For example, a person may decide to use more of his creative and innovative ability as a white-collar worker than as a blue-collar worker.

Based on the model, I identify and estimate the returns to creative and innovative ability, communication ability, and reading and math ability for white-collar and blue-collar workers. A special challenge for the identification and estimation is the partial observation of the supply of ability in work, which is unobserved in occupations that are not chosen. My solution involves projecting the supply of ability in work on the supply in leisure, which is observed for all individuals.

The results show that the individual's supply of ability should be allowed to be endogenous to correct for sample selection bias caused by occupational choice. For white-collar workers, the results from a conventional selection model show that the estimated return to creative and innovative ability is small and statistically insignificant, while the estimated returns to communication ability and reading and math ability are large and statistically significant. Switching to a selection model with endogenous supply of ability, the returns to all three abilities for white-collar workers are estimated to be reasonably large. In this model, the estimated return to creative and innovative ability is now statistically significant, while the estimated returns to communication ability and reading and math ability are imprecisely measured and thus are statistically insignificant. For blue-collar workers, the estimated returns also vary across the two selection models for each ability, but the estimated returns are generally close to zero and statistically insignificant.

These results contribute to the discussion of policies that foster human capital. The insights obtained indicate that society can increase individual abilities in two ways. First, individual abilities can be increased in the traditional way, through education policies. The more that education strengthens individuals' innate abilities, the more that individuals will employ their abilities in work.

Second, individual abilities can be increased through business policies. If politicians implement policies that attract more workers to white-collar work, or blue-collar work, they can affect the total ability supply in the economy. For example, politicians may increase the total supply of creative and innovative ability, communication ability, and reading and math ability if they attract more workers to white-collar work when the returns to the three abilities are higher for white-collar work than for blue-collar work. Thus, a society should consider its business policies in addition to its education policies if it desires to stimulate how much people use their abilities in the workplace.

Appendix A

Derivation of Conditional Expectations

A. Conditional Expectation of the Choice of Occupation

The conditional expectation of the choice of occupation, D, given the observed variables $\varpi = \left\{A_k^{l*}, X, Z\right\}$ equals the probability of the choice of white-collar work, $D = 1$:

$$\mathrm{E}\left[D \mid \varpi\right] = 0\left[1 - \mathrm{P}\left(D = 1 \mid \varpi\right)\right] + 1\left[\mathrm{P}\left(D = 1 \mid \varpi\right)\right] = \mathrm{P}\left(D = 1 \mid \varpi\right) \tag{A1}$$

with the probability being conditional on the observed variables ϖ. The conditional probability can be expressed as

$$\mathrm{P}\left(D = 1 \mid \varpi\right) = \mathrm{P}\left(V \leq g\left(\varpi\right) \mid \varpi\right) = \Phi\left(g\left(\varpi\right)/\sigma_V\right) \tag{A2}$$

given the selection equation that $D = 1$ if $V \leq g\left(\varpi\right)$ and assumptions A.1 and A.2. Inserting (A2) into (A1) gives

$$\mathrm{E}\left[D \mid \varpi\right] = \mathrm{P}\left(D = 1 \mid \varpi\right) = \Phi\left(g\left(\varpi\right)/\sigma_V\right) \tag{A3}$$

This expression is the conditional expectation of the choice of occupation, D, given the observed variables ϖ.

B. Conditional Expectation of the Supply of Ability in Work for White-Collar Workers

The conditional expectation of the supply of ability k in work for white-collar workers, A_{k2}^{w*}, given the observed variables ϖ and the choice of white-collar work, $D=1$, can be expressed as

$$E\left[A_{k2}^{w*} \mid \varpi, D=1\right]=E\left[\theta_{0k2}+\theta_{1k2}A_k^{l*}+\varepsilon_{k2} \mid \varpi, D=1\right] \tag{A4}$$

given the equation for the supply of ability in work for white-collar workers, $A_{k2}^{w*}=\theta_{0k2}+\theta_{1k2}A_k^{l*}+\varepsilon_{k2}$.

Rearranging (A4) yields

$$E\left[A_{k2}^{w*} \mid \varpi, D=1\right]=\theta_{0k2}+\theta_{1k2}A_k^{l*}+E\left[\varepsilon_{k2} \mid \varpi, D=1\right] \tag{A5}$$

It follows from the selection equation that $D=1$ if $V \leq g(\varpi)$. Inserting this expression into (A5) gives

$$E\left[A_{k2}^{w*} \mid \varpi, D=1\right]=\theta_{0k2}+\theta_{1k2}A_k^{l*}+E\left[\varepsilon_{k2} \mid \varpi, V \leq g(\varpi)\right] \tag{A6}$$

The conditional expectation of the unobserved random variable in the equation for the supply of ability in work for white-collar workers, ε_{k2}, given the observed variables ϖ and the truncation, $V \leq g(\varpi)$, is

$$E\left[\varepsilon_{k2} \mid \varpi, V \leq g(\varpi)\right]=E\left[\varepsilon_{k2} \mid V \leq g(\varpi)\right] \tag{A7}$$

from assumption A.1. The conditional expectation of ε_{k2}, given the truncation $V \leq g(\varpi)$, is

$$E\left[\varepsilon_{k2} \mid V \leq g(\varpi)\right]=-\rho_{V,\varepsilon_{k2}}\sigma_{\varepsilon_{k2}}\frac{\varphi\left(g(\varpi)/\sigma_V\right)}{\Phi\left(g(\varpi)/\sigma_V\right)} \tag{A8}$$

from assumption A.3 (see Greene (2003) pp. 781-782).

Inserting (A8) into (A6) gives

$$E\left[A_{k2}^{w*} \mid \varpi, D=1\right]=\theta_{0k2}+\theta_{1k2}A_k^{l*}+\rho_{V,\varepsilon_{k2}}\sigma_{\varepsilon_{k2}}\lambda_2 \tag{A9}$$

where $\lambda_2 = -\varphi\left(g(\varpi)/\sigma_V\right)/\Phi\left(g(\varpi)/\sigma_V\right)$. This expression is the conditional expectation of the supply of ability k in work for white-collar workers, A_{k2}^{w*}, given the observed variables ϖ and the choice of white-collar work, $D=1$.

C. Conditional Expectation of Income for White-Collar Workers

The conditional expectation of income for white-collar workers, Y_2, given the observed variables ϖ and the choice of white-collar work, $D=1$, is

$$E\left[Y_2 \mid \varpi, D=1\right] = E\left[\beta_{02} + \sum \beta_{k2} A_{k2}^{w*} + \beta_{X2} X + U_2^Y \mid \varpi, D=1\right] \tag{A10}$$

given the income equation, $Y_2 = \beta_{02} + \sum \beta_{k2} A_{k2}^{w*} + \beta_{X2} X + U_2^Y$. Rearranging (A10) yields

$$E\left[Y_2 \mid \varpi, D=1\right] = \beta_{02} + \beta_{X2} X + \sum \beta_{k2} E\left[A_{k2}^{w*} \mid \varpi, D=1\right] + E\left[U_2^Y \mid \varpi, D=1\right] \tag{A11}$$

The conditional expectation of the unobserved random variable in the income equation for white-collar workers, U_2^Y, given the observed variables ϖ and the choice of white-collar work, $D=1$, is

$$E\left[U_2^Y \mid \varpi, D=1\right] = -\rho_{V,U_2^Y} \sigma_{U_2^Y} \frac{\varphi\left(g(\varpi)/\sigma_V\right)}{\Phi\left(g(\varpi)/\sigma_V\right)} \tag{A12}$$

from assumptions A.1 and A.2. The derivation follows the steps in (A7) and (A8).

Inserting (A12) and the expression for $E\left[A_{k2}^{w*} \mid \varpi, D=1\right]$ from (A4) into (A11) gives

$$E\left[Y_2 \mid \varpi, D=1\right] = \beta_{02} + \beta_{X2} X + \sum \beta_{k2}\left[\theta_{0k2} + \theta_{1k2} A_k^{l*} + \rho_{V,\varepsilon_{k2}} \sigma_{\varepsilon_{k2}} \lambda_2\right] + \rho_{V,U_2^Y} \sigma_{U_2^Y} \lambda_2 \tag{A13}$$

Equation (A13) simplifies to

$$E\left[Y_2 \mid \varpi, D=1\right] = \gamma_{02} + \sum \gamma_{k2} A_k^{l*} + \beta_{X2} X + \left[\sum \beta_{k2} \rho_{V,\varepsilon_{k2}} \sigma_{\varepsilon_{k2}} + \rho_{V,U_2^Y} \sigma_{U_2^Y}\right] \lambda_2 \tag{A14}$$

where $\gamma_{02} = \beta_{02} + \sum \beta_{k2}\theta_{0k2}$ and $\gamma_{k2} = \beta_{k2}\theta_{1k2}$. This expression is the conditional expectation of the income for white-collar workers, Y_2, given the observed variables ϖ and the choice of white-collar work, $D = 1$.

Appendix B

Survey Questions and Response Scales

A. Survey Questions

The creative and innovative ability, communication ability, and reading and math ability are measured as follows.

The supply of creative and innovative ability in work is determined via three questions: Have you developed or helped to develop new products or services within the last three months at your place of work? Have you participated in testing new methods of working within the last three months at your place of work? Does your job require you to contribute with innovative thinking?

The supply of creative and innovative ability in leisure is measured via one question concerning the creation of knowledge that can be transferred from one context to another: Do you think of ideas during your leisure that could be used in work?

The supply of communication ability in work is measured via three questions concerning the use of communication methods and three questions assessing the use of communication tools: How frequently do you make presentations, give instructions, or the like to a group of people at work? How frequently do you deal with cases or problems that others discuss with you at work? How frequently do you deal with cases or problems that others present to you in writing at work? How often do you write letters/e-mails as part of your job? How often do you search for information on the internet as part of your job? How often do you use an ordinary phone as part of your job?

The supply of communication ability in leisure is measured via five questions regarding the use of communication tools: How often do you do write letters/e-mails in your leisure? How often do you search for information on the internet in your leisure? How often do you use a computer in your leisure? How often do you use an ordinary phone in your leisure? How often do you use a cell phone in your leisure?

The supply of reading and math ability in work is assessed by two questions: How often do you have to read as part of your job? How often do you have to use math or arithmetic in your work?

The supply of reading and math ability in leisure measures reading comprehension via three questions: Have you read one or more books within the last six months in your leisure? How often do you read newspapers, journals, or magazines in your leisure? In your leisure, how often do you write something that fills more than one page? Math comprehension in leisure is not measured.

B. Response Scales

The responses offered for the questions on creative and innovative ability are “to a very large degree,” “to a large degree,” “to some degree,” “to a lesser degree,” and “not at all.” The responses offered for the questions on communication ability and reading and math ability are “every day,” “every week,” “every month,” and “never.” The exception is the question on reading books in leisure, for which the responses are “yes” and “no.”

Appendix C

Index Construction and Estimation Techniques

A. Index Construction

I create one index for each individual ability supplied in work and in leisure. All indices for the abilities, other than two of them, are constructed using factor analysis, as they are measured with three or more underlying variables each. Factors have means equal to zero for the full sample in the summary statistics and means equal to zero for each occupation type in the estimation of the returns.

The two exceptions are the reading and math ability supplied in work and creative and innovative ability supplied in leisure. They are based on simple indices because they are determined by two underlying variables and one underlying variable, respectively. I standardize the underlying variables such that they have means equal to zero and variances equal to one for the full sample in the summary statistics and in the estimation of the returns. The simple index based on two variables gives equal weight to each variable.

B. Estimation Techniques

I estimate the regression and the factor scores simultaneously in the basic wage regression and in each step of the selection model estimations, respectively. See Jöreskog, Sörbom, and Magdison (1979) and Borghans, Duckworth, Heckman, and Weel (2008) for an introduction to factor analysis and Bollen (1989) and Browne and Arminger (1995) for a description of simultaneous estimation. To take into account that the survey data are categorical, I use polychoric correlations between the underlying variables in the estimation of the factor scores.

The bootstrapped standard errors in the selection models are based on 500 samples. The standard error is calculated as the interquartile range divided by 1.34, that is,

$\sigma = (75\text{th} - 25\text{th percentile})/1.34$. The significance levels are based on the assumption that the data are normally distributed with respect to the estimates and standard errors.

References

Bollen, Kenneth A. 1989. *Structural Equations with Latent Variables*. New York: Wiley.

Borghans, Lex, Angela Lee Duckworth, James J. Heckman, and Bas Ter Weel. 2008. "The Economics and Psychology of Personality Traits." *Journal of Human Resources* 43 (4): 972-1059.

Browne, Michael W., and Gerhard Arminger. 1995. "Specification and Estimation of Mean- and Covariance-Structure Models." In *Handbook of Statistical Modeling for the Social and Behavioral Sciences*, edited by Gerhard Arminger, Clifford C. Clogg, and Michael E. Sobel, 311-359. New York: Plenum Press.

Cawley, John, James J. Heckman, and Edward Vytlacil. 1999. "Meritocracy in America: Wages Within and Across Occupations." *Industrial Relations* 38 (3): 250-296.

Cunha, Flavio, and James J. Heckman. 2009. "The Economics and Psychology of Inequality and Human Development." *Journal of the European Economic Association* 7 (2-3): 320-364.

Das, Mitali, Whitney K. Newey, and Francis Vella. 2003. "Nonparametric Estimation of Sample Selection Models." *Review of Economic Studies* 70 (1): 33-58.

Gould, Eric D. 2002. "Rising Wage Inequality, Comparative Advantage, and the Growing Importance of General Skills in the United States." *Journal of Labor Economics* 20 (1): 105-147.

Greene, William H. 2003. *Econometric Analysis*. Upper Saddle River, New Jersey: Prentice Hall.

Hartog, Joop. 1977. "On the Multicapability Theory of Income Distribution." *European Economic Review* 10 (2): 157-171.

Heckman, James J. 1976. "The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models." In *Annals of Economic and Social Measurement* 5 (4): 475-492.

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

Heckman, James J., and Bo E. Honoré. 1990. "The Empirical Content of the Roy Model." *Econometrica* 58 (5): 1121-1149.

Heckman, James. J., Lance J. Lochner, and Petra Todd. 2006. "Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond." In *Handbook of the Economics of Education*, vol. 1, edited by Erik A. Hanushek and Finis Welch, 307-458. Amsterdam: North-Holland.

Heckman, James J., and Guilherme Sedlacek. 1985. "Heterogeneity, Aggregation, and Market Wage Functions: An Empirical Model of Self-Selection in the Labor Market." *Journal of Political Economy* 93 (6): 1077-1125.

Heckman, James J., and Guilherme L. Sedlacek. 1990. "Self-Selection and the Distribution of Hourly Wages." *Journal of Labor Economics* 8 (1), Part 2: Essays in Honor of Albert Rees: 329-363.

Heckman, James. J., Jora Stixrud, and Sergio Urzua. 2006. "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior." *Journal of Labor Economics* 24 (3): 411-482.

Hermann, Stefan (ed.). 2005. *The National Competence Account: Main Report* (Det Nationale Kompetenceregnskab: hovedrapport). Copenhagen: Department of Education.

Jöreskog, Karl G., Dag Sörbom, and Jay Magdison. 1979. *Advances in Factor Analysis and Structural Equation Models*. Cambridge, MA: Abt Books.

Klynge, Alice H. 2011. "The Returns to the Multiple Abilities." PhD dissertation, University of Copenhagen.

Krueger, Alan B., and David Schkade. 2008. "Sorting in the Labor Market: Do Gregarious Workers Flock to Interactive Jobs?" *Journal of Human Resources* 43 (4): 859-883.

Mincer, Jacob A. 1974. *Schooling, Experience, and Earnings*. New York: Columbia University Press.

Nielsen, Helena Skyt, Marianne Simonsen, and Mette Verner. 2004. "Does the Gap in Family-Friendly Policies Drive the Family Gap?" *Scandinavian Journal of Economics* 106 (4): 721-744.

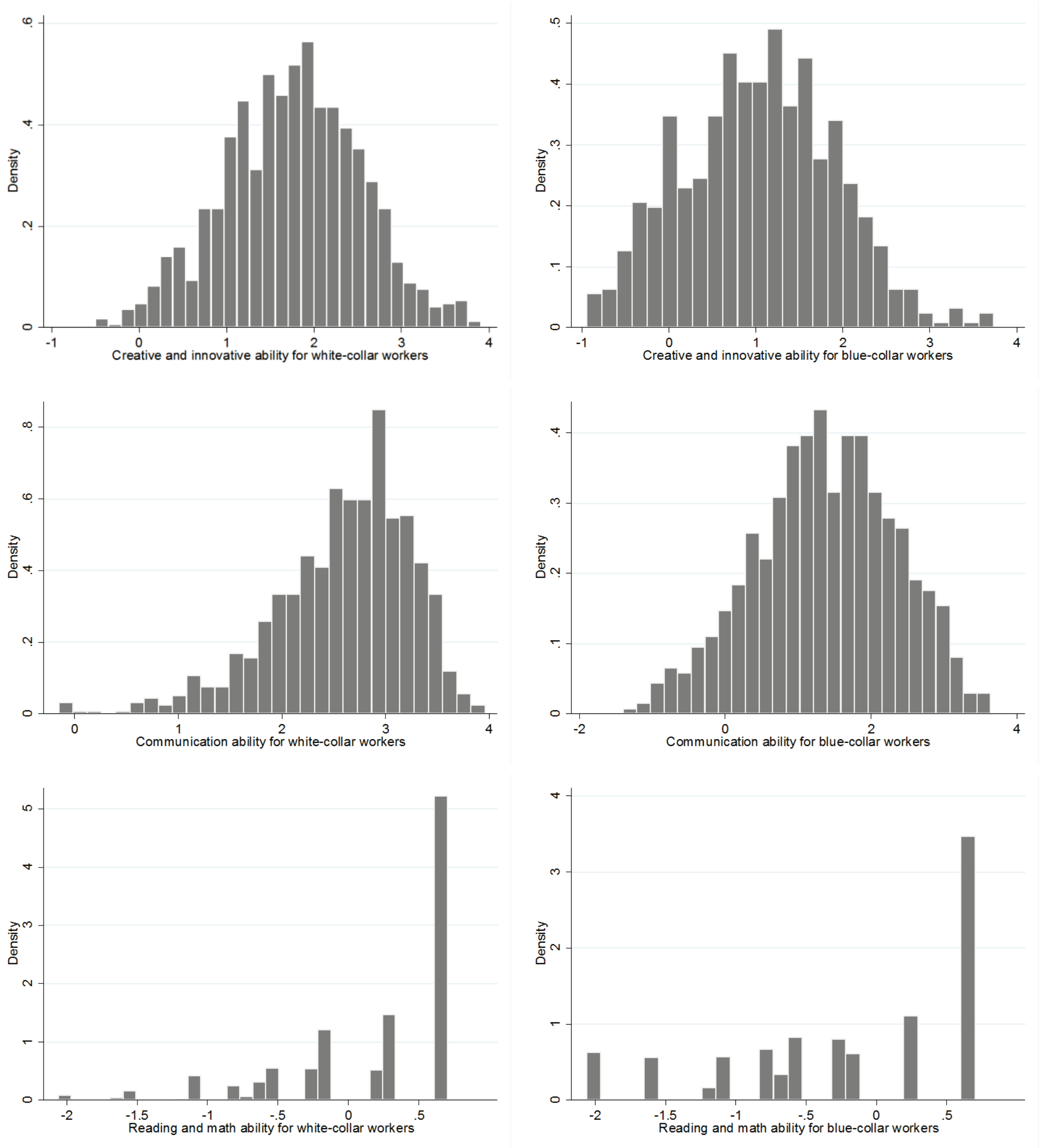
Roy, A. D. 1951. "Some Thoughts on the Distribution of Earnings." *Oxford Economic Papers* 3 (2): 135-146.

Salganik, Laura H., and Dominique S. Rychen (eds.). 2003. *Key Competencies for a Successful Life and a Well-Functioning Society*. Cambridge, MA: Hogrefe & Huber.

Willis, Robert J. 1986. "Wage Determinants: A Survey and Reinterpretation of Human Capital Earnings Functions." In *Handbook of Labor Economics*, vol. 1, edited by Orley Ashenfelter and Richard Layard, 525-602. Amsterdam: North-Holland.

Figure 1

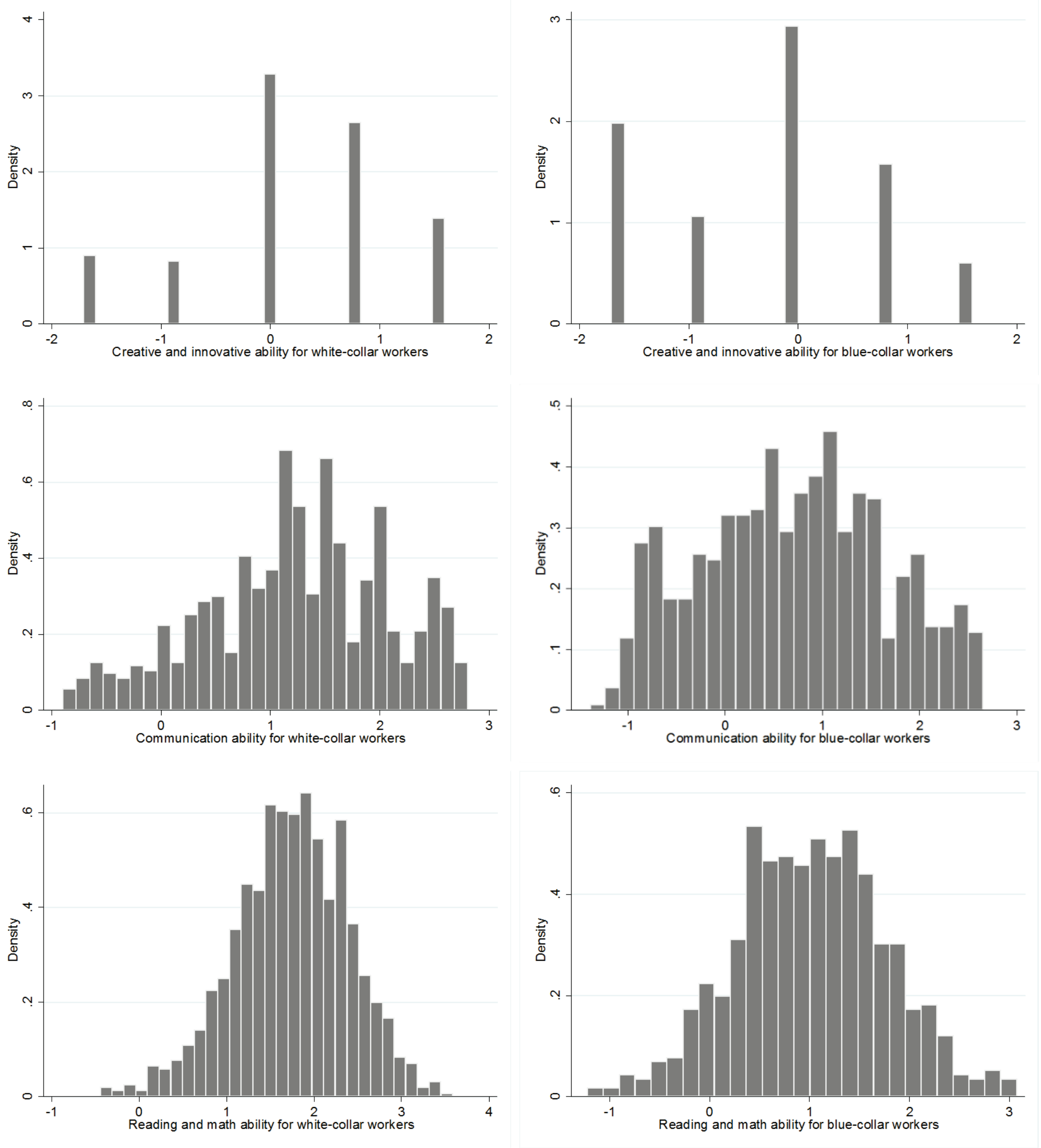
Histograms for Ability Supplied in Work



Note.—The number of white-collar workers is 1,159, and the number of blue-collar workers is 729. The histograms illustrate the index scores for a single ability supplied in work. The indices for creative and innovative ability and communication ability are constructed using factor analysis, while the index for reading and math ability is based on a simple index, measured with two underlying variables only.

Figure 2

Histograms for Ability Supplied in Leisure



Note.—The number of white-collar workers is 1,159, and the number of blue-collar workers is 729. The histograms illustrate the index scores for a single ability supplied in leisure. The indices for communication ability and reading and math ability are constructed using factor analysis, while the index for creative and innovative ability is based on a simple index, measured with one underlying variable only.

Table 1

Means and Correlations for Ability Supplied in Work and in Leisure

Occupation Type	White-Collar Workers	Blue-Collar Workers
Means for ability supplied in work:		
Creative and innovative ability	1.756 (.779)	1.036 (.874)
Communication ability	2.593 (.663)	1.374 (.966)
Reading and math ability	.203 (.622)	-.185 (.870)
Means for ability supplied in leisure:		
Creative and innovative ability	.196 (.944)	-.285 (1.018)
Communication ability	1.230 (.844)	.697 (.945)
Reading and math ability	1.744 (.653)	1.013 (.754)
Correlations between ability supplied in work and in leisure:		
Creative and innovative ability	.463 ***	.434 ***
Communication ability	.453 ***	.433 ***
Reading and math ability	.199 ***	.242 ***
Number of observations	1,159	729

Note.—Standard deviations are in parentheses. Means and correlations are based on index scores for ability supplied.
*** Significant at 1 percent.

Table 2

Means for Outcome and Background Variables

Occupation Type	White-Collar Workers		Blue-Collar Workers	
Wage	203.79	(87.651)	165.80	(44.308)
Education	13.579	(2.562)	10.630	(1.420)
Labor market experience	14.097	(7.489)	14.890	(7.381)
Male	.418	…	.646	…
Father white-collar	.699	…	.502	…
Number of observations	1,159		729	

Note.—Standard deviations are in parentheses.

Table 3

Results of Step 1 for the Selection Models

Dependent variable: Dummy variabel for white-collar or blue-collar		
Model	Selection Model with Endogenous Supply of Ability	Conventional Selection Model
Exclusion restriction		
Father white-collar	.283 ***	.332 ***
	(.077)	(.098)
Ability supplied		
Creative and innovative ability	.187 ***	-.135 **
	(.038)	(.064)
Communication ability	.183 ***	.897 ***
	(.063)	(.071)
Reading and math ability	.195 **	.537 ***
	(.084)	(.067)
Control		
Education	.395 ***	.487 ***
	(.021)	(.028)
Labor market experience	-.003	-.013
	(.020)	(.025)
Experience-squared	.080	.123
	(.062)	(.077)
Male	-.708 ***	-.976 ***
	(.077)	(.100)
Constant	4.368 ***	5.264 ***
	(.296)	(.383)
Number of observations	1,888	1,888

Note.—Standard errors are in parentheses. Ability supplied is ability supplied in leisure in the selection model with endogenous supply of ability and ability supplied in work in the conventional selection model due to the models' setup.

** Significant at 5 percent.

*** Significant at 1 percent.

Table 4

Results of Step 2 for the Selection Model with Endogenous Supply of Ability

Occupation Type	White-Collar Workers	Blue-Collar Workers
Dependent variable: Creative and innovative ability supplied in work		
Creative and innovative ability supplied in leisure	.725 ***	.629 ***
	(.071)	(.077)
Lambda	.636 ***	.245 **
	(.128)	(.106)
Constant	…	…
	…	…
Dependent variable: Communication ability supplied in work		
Communication ability supplied in leisure	.405 ***	.243 ***
	(.046)	(.053)
Lambda	.733 ***	.926 ***
	(.114)	(.092)
Constant	…	…
	…	…
Dependent variable: Reading and math ability supplied in work		
Reading and math ability supplied in leisure	.081 **	.081
	(.035)	(.060)
Lambda	.028	.262 ***
	(.057)	(.066)
Constant	.211 ***	-.393 ***
	(.023)	(.065)
Number of observations	1,159	729

Note.—Standard errors are in parentheses. Standard errors are bootstrapped. There is no estimate for the constant for creative and innovative ability and communication ability because their indices for supply of ability in work are based on factor analysis, which implies that that the intercept is zero within each occupation.

** Significant at 5 percent.

*** Significant at 1 percent.

Table 5

Results of Step 3 for the Selection Model with Endogenous Supply of Ability

Dependent variable: Log(hourly wage)		
Occupation Type	White-Collar Workers	Blue-Collar Workers
Ability supplied		
Creative and innovative ability	.041 ***	.001
	(.010)	(.010)
Communication ability	.023	.000
	(.016)	(.012)
Reading and math ability	.006	-.010
	(.020)	(.016)
Control		
Education	.030 ***	-.018 *
	(.007)	(.010)
Labor market experience	.018 ***	.019 ***
	(.004)	(.005)
Experience-squared	-.026 **	-.049 ***
	(.013)	(.015)
Male	.281 ***	.273 ***
	(.022)	(.024)
Selection correction		
Lambda	.083 *	.092 **
	(.049)	(.038)
Constant	4.553 ***	4.868 ***
	(.110)	(.091)
Number of observations	1,159	729

Note.—Standard errors are in parentheses. Standard errors are bootstrapped. Ability supplied is ability supplied in leisure due to the model's setup.

* Significant at 10 percent.
** Significant at 5 percent.
*** Significant at 1 percent.

Table 6

Returns to Abilities for White-Collar Workers

Dependent Variable: Log(hourly wage)			
Model	Basic Wage Regression	Conventional Selection Model	Selection Model with Endogenous Supply of Ability
Ability supplied			
Creative and innovative ability	.007	.002	.056 ***
	(.011)	(.011)	(.013)
Communication ability	.121 ***	.124 ***	.056
	(.009)	(.013)	(.034)
Reading and math ability	.102 ***	.086 ***	.077
	(.016)	(.013)	(.260)
Control			
Education	.043 ***	.038 ***	.030 ***
	(.003)	(.004)	(.007)
Labor market experience	.016 ***	.015 ***	.018 ***
	(.005)	(.004)	(.004)
Experience-squared	-.017	-.017	-.026 **
	(.015)	(.013)	(.013)
Male	.255 ***	.267 ***	.281 ***
	(.017)	(.020)	(.022)
Selection correction			
Lambda	…	.148 ***	.083 *
	…	(.053)	(.049)
Decomposition of selection effect			
Total ability supply selection effect	…	…	.079 ***
	…	…	(.026)
Creative and innovative ability selection effect	…	…	.036 ***
	…	…	(.010)
Communication ability selection effect	…	…	.041 **
	…	…	(.018)
Reading and math ability selection effect	…	…	.002
	…	…	(.009)
Conventional selection effect	…	…	.004
	…	…	(.056)
Constant	4.362 ***	4.438 ***	…
	(.051)	(.068)	…

Note.—Number of observations is 1,159. Standard errors are in parentheses. Standard errors are bootstrapped in the selection models.

* Significant at 10 percent; ** significant at 5 percent; *** significant at 1 percent.

Table 7

Returns to Abilities for Blue-Collar Workers

Dependent Variable: Log(hourly wage)

Model	Basic Wage Regression	Conventional Selection Model	Selection Model with Endogenous Supply of Ability
Ability supplied			
Creative and innovative ability	.010	.008	.001
	(.014)	(.015)	(.015)
Communication ability	.008	.003	.001
	(.013)	(.015)	(.030)
Reading and math ability	.015	.001	-.118
	(.010)	(.012)	(.146)
Control			
Education	.004	-.007	-.018 *
	(.006)	(.009)	(.010)
Labor market experience	.018 ***	.019 ***	.019 ***
	(.005)	(.005)	(.005)
Experience-squared	-.042 ***	-.046 ***	-.049 ***
	(.014)	(.015)	(.015)
Male	.235 ***	.252 ***	.273 ***
	(.018)	(.021)	(.024)
Selection correction			
Lambda	…	.031 *	.092 **
	…	(.018)	(.038)
Decomposition of selection effect			
Total ability supply selection effect	…	…	-.030
	…	…	(.020)
Creative and innovative ability selection effect	…	…	.000
	…	…	(.002)
Communication ability selection effect	…	…	.001
	…	…	(.019)
Reading and math ability selection effect	…	…	-.031
	…	…	(.022)
Conventional selection effect	…	…	.122 **
	…	…	(.051)
Constant	4.733 ***	4.805 ***	…
	(.077)	(.086)	…

Note.—Number of observations is 729. Standard errors are in parentheses. Standard errors are bootstrapped in the selection models.

* Significant at 10 percent; ** significant at 5 percent; *** significant at 1 percent.