

Review of Philipp KRAFT's dissertation

Three Essays on the Economics of Beauty

Olivier GERGAUD,

Professor of Economics,

Bordeaux Management School and University of Reims

I really enjoyed reading this thesis which is based on three original contributions to the burgeoning field called "Economics of beauty" mostly interested in assessing the various economic impacts beauty may have on some key labor market outcomes like employment and salaries.

Comments regarding the first paper:

The first paper is an attempt to measure the influence of physical appearance (face) on the rate at (and speed with) which virtual students, close to graduation, receive positive callbacks after spontaneous applications. The literature review is short, focused and based on the most relevant papers coming from different fields of research. Both the research strategy and its limitations are carefully stated. The field experiment is well designed and the framework is introduced very clearly.

I have more concerns about the econometrics of the paper which could be easily improved. First, I would recommend using a probit or logit analysis instead of a -less appropriate- linear probability model to regress the probability with which these virtual guinea pigs receive a positive callback. Second, I would try a more comprehensive regression that would include monthly time dummies to control for some potential seasonal effects as the survey took place over several months. I would also control for gender differences using a dummy. I would be curious to check the level of determination of the regression which is not provided here (Pseudo R-square). Why didn't you introduce the treatment (beauty) as a rating instead of a basic dummy? Are your results different if beauty is controlled for using an ordinal or even better as a continuous rating? Third, why wouldn't you regress time elapsed between the application and the response? This would improve the analysis and maybe we would observe here a so-called beauty-hungry effect on this other aspect of discrimination. I suggest using a discrete survival analysis (probit) to achieve this goal. Did you test for the presence of outlying values in the sample? Robust logit estimators are now readily available in SPLUS and R¹ and applying these techniques often leads to dramatically different conclusions. See Croux and Haesbroeck (2003) for more information about these estimators.

Last, I cast serious doubts about the possibility to generalize this analysis to draw general conclusions regarding the entire labor market. Indeed the analysis only concerns very young and rather inexperienced workers and employers' expectations regarding a sample of interns is likely lower than

¹ See <http://www.econ.kuleuven.be/public/NDBAE06/programs/#roblog> for more information about these commands.

with more experienced and permanent workers. The economic consequences of a bad choice here are not as high compared to these of a bad choice with a permanent worker. I would appreciate to discuss this point with Philipp over the defense.

Comments regarding the second paper:

The second paper is an attempt to capture the influence of intelligence, education and beauty on wages using original data from the National University of La Plata in Argentina.

Intelligence is an objective measure (IQ score) here. It is unclear how Beauty was assessed at first sight through the mean of a series of pictures. Very little information is provided about the rating process, the type and quality of pictures used in particular. Are these standardized pictures? This is a very important point as pictures of heterogeneous quality could introduce a bias in the empirical analysis.

I don't get the intuition behind the selection process here (page 50). Why spousal wage and number of children would be valid instruments that would identify 64 very different individuals in the sample? Indeed, these people either work (25 ind.) or are unemployed (39 ind.). The author claims that these instruments are good predictors of the probability of working. But unemployed people are not working by definition contrarily to the other 25 people who are working. Are these unemployed people inactive? Are they looking for a job? This identification strategy is confusing to me. My intuition is that you shouldn't merge these categories identifying very different people. The first strategy could be to work with two participation equations instead of just one with appropriate instruments for each. Another strategy would be to reduce the sample and ignore one of these two categories or people. The results for the usual tests (overidentification, exogeneity and weak instruments tests) are not reproduced here either (see section 3.4). A more conventional approach should consist in relying on the Stock and Yogo (2005) statistics to test for the weakness of the instruments instead of an informal rule of thumb based on the F-statistic. The results for the Sargan–Hansen test of overidentifying restrictions should also be reproduced to better convince the reader about the validity of the instruments.

My major concern with this part of the thesis is that the endogeneity of education with respect to Intelligence is not treated here. A substantial fraction of the literature on wage equations deals with this endogeneity issue and the necessity to find valid instruments to get an unbiased estimate of the impact of education on earnings. This part of the thesis would require some other references to the pioneering works of Orley Ashenfelter, David Card, Alan Krueger, Sherwin Rosen and Cecilia Rouse among others. In this literature "parental education" is often used as an instrument for the level of education and the appropriate estimation procedure is 2SLS and not just LS (Heckman here). In a recent paper, Gergaud and Verardi (2007) combine the Heckman selection model with a 2SLS approach (see the paper for further details about this procedure that could easily be applied to this specific framework).

At the same time I'm concerned with another endogeneity issue, that of education (proxy for productivity) or Intelligence with respect to Beauty and experience. This would require additional statements to convince the reader (relying on the usual tests provided by e.g. `ivreg2`) that these

variables are exogenous variables. I would also appreciate to take a close look at the correlation matrix. That would help the reader to get a precise idea about the evolution of beauty ratings over time (correlation with age which is used as a proxy for experience here) but more generally to gauge the correlation between the main variables of interest that are Beauty, Intelligence, Education and Experience.

I enjoyed reading the section about the results for the series of tenure-specific regressions. This type of analysis is particularly insightful and promising. Overall, it seems to me that this paper is the most promising one even if it requires substantial additional efforts to get rid of the series of econometric issues raised above to be able to make it in a good journal in Labor Economics. This project is highly stimulating and I do believe that Philipp Kraft will be able to improve the analysis.

Comments regarding the third paper:

The third paper tests whether there is a beauty premium for CEOs. This original study is based on a large sample of 450 German CEOs and the results tend to show that these enjoy a sizeable beauty premium. The first added value that I see with this contribution is the dataset that has been created by Philipp Kraft for this study.

Pictures used to assess beauty come from the CEO's CV that is available on the company's website. I'm afraid one more time that these pictures are not of homogeneous quality. This would mean that beauty is observed here with error and that some CEOs look better on the picture than in reality thanks to a more efficient photographer (while the opposite is true). This is not a major concern as all of these firms certainly work with competent photographers but still. This concern and limitation of the study should be clearly indicated somewhere in the text.

The idea of using a control group of 25 Austrian people to prevent some sort of proximity or familiarity bias is a very good point.

Equation 1 should be rewritten with different subscripts to make it clear that the log of earnings of CEO i at time t ($\log(\text{comp}_{it})$) depends on information about the performance of the firm lagged (comp_{it-1})... ε_{it} should replace ε_i .

The level of education is proxied by the number of years of schooling. I'm assuming that these CVs contain better information about education than just a quantitative measure. I would introduce some qualitative variables such as the type of diploma, the reputation of the school or university where the CEO graduated and so on. Same remark with experience, a measure that could be easily refined with information about the type of experience such as the size of the previous companies run by the CEO.

Why do you need to standardize the beauty index? Why don't you use an ordinal scale or a set of dummies instead? That would make the interpretation easier.

I disagree with the interpretation of the coefficient of beauty which is said to be highly significant while significantly different from 0 at the 5% level. That is a bit exaggerated. Moreover, the result on the impact of beauty, which is the focus of the paper, seems to be highly sensitive to the type of

specification used. Indeed the coefficient of Beauty is no longer significant in the second specification when the author introduces a series of interaction terms.

Table 3 presents rather weak evidence of a decline of the influence of beauty on earnings with border significant coefficients even if I confess that we seem to observe a decline over the tenure as well as with experience. I appreciate the robustness check section which raises a series of quite relevant points.

As a consequence I would be very careful with the conclusions drawn from this analysis based on estimations that clearly need to be refined. This paper seems to be the least accomplished of the three contributions. It however has a good publication potential.

Minor comment: you should tell earlier in the text (page 87) that the beauty index has been standardized.

General conclusions:

Once again, I really enjoyed reviewing this work which seems important to me. I have good hope that Philipp Kraft will publish these essays in respectable journals. The author however needs to clarify several points, to strengthen the econometric analysis in several ways.

References :

Croux, C., and Haesbroeck, G. (2003), "Implementing the Bianco and Yohai estimator for logistic regression", *Computational Statistics and Data Analysis*, 44, 273-295.

Gergaud, O., Montana, L. and V. Verardi, (2007), "Stardust over Paris Gastronomic Restaurants", *Journal of Wine Economics*, 2(1), 24-39.

Stock, J.H. and Yogo, M. (2005), "Testing for Weak Instruments in Linear IV Regression". In D.W.K. Andrews and J.H. Stock, eds. *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*. Cambridge: Cambridge U. Press, pp. 80–108. Working paper version: NBER Technical Working Paper 284. <http://www.nber.org/papers/T0284>.