Charles University

Faculty of Social Sciences

Institute of Economic Studies

MASTER'S THESIS

Beauty and Productivity: A Meta-Analysis

Author: **Ing. Kseniya Bortnikova**

Supervisor: **Doc. Tomas Havranek, Ph.D.**

Academic Year: **2018/2019**

Declaration of Authorship

The author hereby declares that he compiled this thesis independently, using only the listed resources and literature, and the thesis has not been used to obtain a different or the same degree.

The author grants to Charles University permission to reproduce and to distribute copies of this thesis document in whole or in part.

Prague, May 10, 2019

Signature

Acknowledgments

This thesis is part of a project that has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No. 681228.

I would like to express my gratitude to doc. PhDr Tomás Havránek, Ph.D., for his patient guidance, valuable advice and especially, for encouraging words. I am eternally grateful to my family for motivating me when it was most needed. I would also like to thank my friends, Olesia and Sasha, for their support.

Abstract

This thesis conducts a quantitative synthesis of 418 estimates of the effect of beauty on productivity as reported in 37 studies. We test the estimates of beauty effect for publication selection, using informal testing of the funnel plot as well as formal testing methods. We find solid evidence of selective reporting: positive estimates of the beauty effect are preferred in literature. To determine the sources of heterogeneity in the reported estimates, we collect the set of 21 explanatory variables. We take the model uncertainty into account and employ the Bayesian model averaging; the Frequentist model averaging is used as a robustness check. The results indicate that differences in the reported estimates appear to be driven by choice of study design and sources of real heterogeneity, such as geographical regions and individual characteristics of respondents (age, education and cognitive skills). The type of occupation and gender of respondents have no impact on the estimates of beauty effect in relation to productivity. The average beauty effect is probably much lower than commonly believed based on the available empirical literature.

JEL Classification	C83, J3,J7, M51
Keywords	meta-analysis, beauty bias, productivity, discrimination, publication bias

Author's e-mail	xenia.bortnikova@gmail.com
Supervisor's e-mail	tomas.havranek@fsv.cuni.cz

Abstrakt

Předložená práce provádí kvantitativní syntézu 418 odhadů efektu atraktivity na produktivitu, které byly získány ze 37 studií. Za účelem zjištění míry publikační selektivity jsem tyto odhady efektu atraktivity podrobila testům prostřednictvím neformálních testovacích metod jako je trychtýřový graf, jakož i prostřednictvím metod formálních. Pomocí nich jsem dospěla k přesvědčivým důkazům o existující selektivitě v uváděných údajích: V literatuře jednoznačně převažují pozitivní odhady efektu atraktivity. K určení zdrojů heterogenity v uvedených odhadech jsem shromáždila 21 vysvětlujících proměnných. V práci zohledňuji nejistotu modelů a spoléhám se na Bayesovské průměrování modelů. Pro potřeby kontroly robustnosti používám frekventistické průměrování modelů. Výsledky zkoumání naznačují, že rozdíly v uváděných odhadech jsou zřejmě důsledkem odlišné volby koncepce a různých zdrojů pravé heterogenity, jako jsou zeměpisné regiony a individuální rysy respondentů (věk, vzdělání a kognitivní schopnosti).Povolání a pohlaví respondentů nemají žádný vliv na odhady dopadů efektu atraktivity na produktivitu. Průměrná míra efektu atraktivity je pravděpodobně mnohem nižší než se mnozí domnívají na základě dostupné empirické literatury.

JEL kódy:	C83, J3, J7,M51
Klíčová slova:	meta-analýza,produktivita,diskriminace,publikační selektivita

E-mail autora	`xenia.bortnikova@gmail.com`
E-mail vedoucího práce	`tomas.havranek@fsv.cuni.cz`

Contents

Acronyms

BE Between Effects Model

BMA Bayesian Model Averaging

FAT Funnel Assymetry Test

FMA Frequentist Model Averaging

GMM Generalized Method of Moments

IV Instrumental Variable

MRA Meta-Regression Analysis

OLS Ordinary Least Squares

PET Precision Effect Test

PCC Partial Correlation Coefficient

PIP Posterior Inclusion Probability

PMP Posterior Model Probability

UIP Unit Information Prior

WLS Weighted Least Squares

List of Figures

List of Tables

Master's Thesis Proposal

Author	Ing. Kseniya Bortnikova
Supervisor	Doc. Tomas Havranek, Ph.D.
Proposed topic	Beauty and Productivity: A Meta-Analysis

Motivation Recent social sciences literature shows the existence of physical attractiveness bias across all types of industries.The effect of bias is that good-looking people are treated more positively than those who are less physically attractive.

Researchers in the field of economics repeatedly found a strong impact of physical attractiveness on labour market: people who are assessed as attractive have a greater chance of finding employment and earning more. The effect arises from discrimination or it can reflect an association between attractiveness and productivity. Harper (2000) used British Cohort data and found that both occupation-specific discrimination and productivity effects arising from customer discrimination are significant sources of bias. Distinguishing effects resulting from productivity differences and those resulting from discrimination is complicated by the fact that examined groups may have different labor productivity.

Beauty bias has been observed in many different professions. A significant number of studies examined occupations for which attractiveness is likely to play an important role such as lawyers (Biddle and Hamermesh, 1998), politics (Hamermesh, 2006),restaurant servers (Parrett, 2015). Several studies found evidence of physical attractiveness bias for occupations which do not require face-to-face interaction. Paphawasit and Fidrmuc (2015) found significantly positive effect of authors? attractiveness on research productivity in academic writing. A number of academic papers reported reverse, so-called "beauty is beastly" effect. The effect suggests that attractiveness can be disadvantageous in certain employment context for women (Heilman and Saruwatari, 1979; Johnson, 2010).There are many other factors such as gender, geographical location, cultural characteristics that need to be considered when examining the effect of attractiveness on productivity.

The meta-analytic approach which was introduced by Stanley and Jarrell (1989) is appropriate for providing quantitative review of selected findings in the impact of physical attractiveness on productivity literature.There are number of meta-analytic reviews of studies which examine bias effect of physical attractiveness on a variety of job-related outcomes to date. Hosoda et al.(2003) reported weighted mean beauty bias effect size of 0.37. The main findings are that beauty bias effect is as important for men as for women and it is decreasing in time. Quite on the contrary Jackson et al.(1995) found that beauty bias has stronger effects on occupational domain of men than of women.

The publication bias problem has not been addressed in meta-analysis of beauty and productivity so far. The effect of publication bias is that researchers prefer not to report statistically insignificant results (Stanley, 2005). The presence of publication bias then affects meta-analysis results considerably.

Hypotheses

Hypothesis #1: The estimated beauty bias effect is affected by publication bias in literature.

Hypothesis #2:The effect of beauty bias varies across different occupations.

Hypothesis #3:The magnitude of publication bias has been decreasing over time.

Methodology At the first step of meta-analysis I will conduct a search of relevant academic articles on beauty and productivity. I will use RePEc, Google Scholar, and Scopus databases and keywords such as "beauty bias", and "physical attractiveness",combined with such keywords as "productivity", "performance evaluation". I will check journal articles and working papers used by recent meta-analytical reviews (Hosoda et al., 2005; Langlois et al., 2000) examining the beauty bias effect on labor-market outcomes.

Following previous meta-analytic reviews,I will formulate decision rules for set of explanatory variables for coding.The decision rules need to be based on theories which are considered applicable for beauty bias studies: implicit personality theory and lack-of-fit model (Heilman,1983).

Computing difference between the means of physically attractive and less attractive groups divided by the relevant denominator for the effect size estimate, I will derive standard effect size estimate. The sign of the differences between means is expected to be positive when good-looking person is rated positively on productivity in comparison with less attractive person. The negative sign is expected when less attractive person is rated positively in comparison with attractive one. Following the standard meta-analytical methodology unweighted mean effect size estimates and sample size weighted mean effect size estimates with respect to homogeneity statistics calculation will be obtained.

Under the assumption of no publication bias, the estimates of weighted mean effect of beauty bias on productivity will be randomly distributed around the true mean effect. if some estimates ends up in selection biased category, the reported estimates will be correlated with standard errors.To dedicate and estimate beauty bias effect on productivity under assumption of presence of publication selection bias in literature I will use of the multilevel random effects model that is found to be robust against the publication selection (Stanley, 2008).

Assuming validity of "economic-research-cycle" hypothesis I expect the value of publication bias might have decreased over time. To test the last thesis hypothesis it will be necessary to add an interaction term between the year of publication of the study and the reported standard error.

Expected Contribution Publication bias in beauty and productivity literature has not been addressed so far. The corrected estimates of beauty bias effect are expected to differ significantly from previously reported values from meta-analytical reviews. The expected results of meta-analytic review which will introduce selection bias problem can improve precision of estimates of beauty bias effect on productivity. Updated meta-analysis will offer more systematic and unbiased view at empirical studies.

Outline

1. Introduction section of the thesis will introduce the idea of the topic and will explain the motivation.

2. Literature review section will provide discussion on how different authors examine the beauty bias effect on productivity in the literature.

3. Meta-analysis methodology and data. This section will describe how beauty bias effect size and standard errors have been collected from studies. It will present the methodology of publication selection bias and heterogeneity identification and estimation.

4. Empirical results section will provide the main findings and interpretation of the results.

5. Concluding remarks part will summarise all the findings, the policy implication will be discussed in this section.

Core bibliography

1. Biddle, J. and Hamermesh, D. (1998). ?Beauty, productivity, and discrimination: Lawyers' looks and lucre?. Journal of Labor Economics, 16(1), pp.172-201.

2. Cipriani, G. and Zago, A. (2011). Productivity or Discrimination? Beauty and the Exams. Oxford Bulletin of Economics and Statistics, 73(3), pp.428-447.

3. Deryugina, T. and Shurchkov, O. (2015). Now you see it, now you don?t: The vanishing beauty premium. Journal of Economic Behavior and Organization, 116, pp.331-345.

4. French, M. (2002). Physical appearance and earnings: further evidence. Applied Economics, 34(5), pp. 569-572.

5. Johnson, S., Podratz, K., Dipboye, R. and Gibbons, E. (2010). Physical Attractiveness Biases in Ratings of Employment Suitability: Tracking Down the "Beauty is Beastly" Effect. The Journal of Social Psychology, 150(3), pp.301-318.

6. Hamermesh, D. (2006). Changing looks and changing discrimination: The beauty of economists. Economics Letters, 93(3), pp.405-412.

7. Harper, B. (2000). Beauty, Stature and the Labour Market: A British Cohort Study. Oxford Bulletin of Economics and Statistics, 62(s1), pp.771-800.

8. Hosoda, M., Stone-Romero, E. and Coats, G. (2003). The Effects Of Physical Attractiveness on Job-related Outcomes: a Meta-Analysis of Experimental Studies. Personnel Psychology, 56(2), pp.431-462.

9. Heilman, M. and Saruwatari, L. (1979). When beauty is beastly: The effects of appearance and sex on evaluations of job applicants for managerial and nonmanagerial jobs. Organizational Behavior and Human Performance, 23(3), pp.360-372.

10. Kraft, P. (2012) The role of beauty in the labor market. Prague: CERGE.

11. Langlois, J., Kalakanis, L., Rubenstein, A., Larson, A., Hallam, M. and Smoot, M. (2000). Maxims or myths of beauty? A meta-analytic and theoretical review. Psychological Bulletin, 126(3), pp.390-423.

12. Paphawasit, B. and Fidrmuc, J. (2017). Beautiful Minds: The Impact of Beauty on Research Productivity in Economics. Working Paper, 17-12.

13. Parrett, M. (2015). Beauty and the feast: Examining the effect of beauty on earnings using restaurant tipping data. Journal of Economic Psychology, 49, pp.34-46.

14. Stanley, T. (2005). Beyond Publication Bias. Journal of Economic Surveys, 19(3), pp.309-345.

15. Stanley, T., Doucouliagos, C. and Jarrell, S. (2008). Meta-regression analysis as the socio-economics of economics research. The Journal of Socio-Economics, 37(1), pp.276-292.

16. Stanley, T. and Jarrell, S. (2005). Meta-Regression Analysis: A Quantitative Method of Literature Surveys. Journal of Economic Surveys, 19(3), pp.299-308.

17. Stanley, T.D., Jarrell, S.B., 1989. Meta-regression analysis: a quantitative method of literature surveys. J. Econ. Survey. 3 (2), 161?170.

Author Supervisor

Chapter 1

Introduction

The beauty bias phenomenon has been discussed among sociologists and economists for the last 50 years. It describes the situation in which physically attractive individuals are treated more positively than those who are seemingly less attractive as it is assumed that "what is beautiful is good".

The beauty effect is often studied in the context of discrimination against the group of unattractive. However, employer or customer discrimination cannot be easily distinguished from real differences in productivity. Understanding the channels through which beauty affects productivity-related outcomes is crucial for policymakers to achieve their goals.

Economists have repeatedly found an impact of physical attractiveness on the labor market: good-looking individuals have a greater chance to be employed, work more productively and receive higher wages. Despite decades of research, consensus on the magnitude of the effect of beauty has not been reached and neither is there agreement on mechanisms through which beauty affects labor outcomes.

This thesis aims to review the empirical literature quantitatively, focusing on the following questions: (1) Does the publication bias affect the estimated beauty effect on productivity in literature? (2) Which factors govern the differences in the results of beauty effect estimates? (3) Is the beauty effect consistent across different types of occupations? To my knowledge, an extensive meta-analysis of the relation between beauty and productivity has not yet been conducted.

In order to address these questions, modern meta-analysis techniques have been applied. The presence of publication selection was tested both visually (using funnel plot) and formally (using funnel tests and alternative approaches). Focusing on the aspects related to data specifications, characteristics, and methodologies, a set of 21 explanatory variables was collected. To address the model uncertainty, the Bayesian model averaging technique was employed followed by a frequentist check of the variables with the highest posterior inclusion probability. Furthermore, the robustness check is conducted using the Frequentist model averaging methodology.

The remainder of this thesis is structured as follows. In Section 2, the literature review is provided and followed by a discussion on how researchers measure beauty and productivity. In Section 3, the methodology description is given; this section also outlines the data collection process and provides a summary statistics of the data set. In Section 4, the presence of publication bias in the literature is tested. Section 4 also focuses on explaining the heterogeneity between the beauty effect estimates. Section 5 concludes this thesis, and the Appendix section provides additional important plots and tables.

Chapter 2

Literature Review

The meta-analysis approach is appropriate for providing a quantitative review of selected findings from the labor economics literature. A meta-analysis, however, cannot completely substitute a traditional literature review. Therefore, the review of the most important works, considering the beauty effect on productivity is provided below.

The effect of beauty has generated a considerable number of studies in the last 50 years. Researches have attempted to answer for such questions as: (1) Is it appropriate to consider physical attractiveness as a productive factor? (2) What is the magnitude of the physical attractiveness effect on productivity? (3) Does the effect of beauty depend on gender and varies across different occupations?

The physical attractiveness stereotype, which implies that beauty is rewarded by higher earnings and higher performance ratings, was initially studied in the context of social psychology. Attractive individuals were expected to be more intelligent and sociable than less attractive individuals. The meta-analysis by Langlois *et al.* (2000) reviewed 1800 empirical studies on beauty and included 919 works in the published study. Further studies by psychologists (Kanazawa & Kovar (2004), Ravina (2008)) show that beauty might be correlated with intelligence, education level, and organizational skills.

The evidence from studies involving experimental economic games confirms that more attractive individuals are expected to be more collaborative and trustworthy than those who are less physically attractive (Wilson & Eckel (2006), Andreoni & Petrie (2008)). Taking into account the findings from psychology, it was expected that beauty can have a positive return in the labor market. The lack of data on physical attractiveness supplemented with economic characteristics, however, might be a reason for rather less attention to beauty impact in the economic literature.

and labor market outcomes was presented by Hamermesh & Biddle (1993). The authors found that above-average looking workers earn 5 percent more per hour than average-looking workers, whereas below-average looking workers earn 9 percent less per hour than their average-looking colleagues. The modeling approach formulated by Hamermesh and Biddle provides a strong foundation for research to date. The proposed wage model is based on the assumption that workers receive a reward from a combination of productivity-related characteristics with a level of physical attractiveness. Depending on a type of occupation, these characteristics may receive different levels of importance. Hence, good-looking workers may choose an occupation that attached low importance on attractiveness even if their productivity-related characteristics enable higher wages to be paid. Assuming that productivity characteristics are non-related with beauty, good-looking workers may earn more than their less attractive colleagues.

Another major research was conducted by Harper (2000). The author examined how physical attractiveness influences labor market outcomes using British longitudinal data, which covers 11407 individuals born in Britain in 1958. The author claims that physical attractiveness has a substantial effect on earnings and employment patterns irrespective of gender.

Further empirical results mostly support the hypothesis that physically attractive people are more likely to be employed, work more productively and receive higher wages than less attractive ones (Harper (2000); Ahn & Lee (2013); Paphawasit and Fidrmuc (2017)). A number of studies, however, have reported that the effect of beauty is overestimated or it is highly context-dependent (Deryugina & Shurchkov (2015); Hernandez-Julian & Peters (2017)).

Some authors reported the reverse, so-called "beauty is beastly" effect, which suggests that attractiveness can be disadvantageous in a certain employment context. In particular, the effect was reported for female applicants for traditionally masculine occupations: the good-looking females were considered less suitable for the position than less attractive ones (Heilman & Saruwatari (1979) , Johnson *et al.* (2010)). According to the findings of Johnson *et al.* (2010), the "beauty is beastly" effect is gender-specific. The authors asked participants to match the photos of applicants with the job descriptions. The results have shown that attractiveness is an advantage only for female candidates who applied for traditionally female jobs. The attractive men were matched with all sorts of jobs whereas the good-looking women were matched mostly with jobs of secretaries and were not matched with traditionally male-dominated positions (e.g., mechanical engineer, security director, research and development manager).

The substantial number of studies indicated that the physical attractiveness of workers is positively correlated with their earnings and performance ratings, but the source of pay differentials remains an open issue. What are the channels through which can beauty influence wages and performance? The researchers mainly agreed that there are two underlying effects of physical attractiveness. The first effect relates to discrimination, which means that employers and consumers prefer working with attractive-looking staff members or may wrongly perceive workers as more capable. The second effect is productivity-enhancing, which may occur when physical attractiveness is a direct determinant of worker's productivity.

Mobius & Rosenblat (2006) analyzed the data from the experimental labor market and had obtained the evidence of discrimination: the good-looking workers received the attractiveness premium, which was not directly related to their actual job performances. The authors argue that the high level of self-confidence and better social skills of good-looking workers are the main sources of a beauty premium. Scholz & Sicinski (2015) have reached similar conclusions on the source of beauty premium: good-looking workers are more confident and have better communication skills, which is translated into higher wages.

Pfann *et al.* (2000) provides evidence of the productivity-enhancing effect of beauty. The authors found that companies employing better-looking executives have higher revenues and hence pay them more. Using the Dutch advertising industry sample the authors have shown that the executive's attractiveness increases the company's revenues to the value, which is significantly higher than the premiums that executives receive. Hence, physical attractiveness enhances the human capital of the companies.

The existence of a positive effect of beauty on an individual's labor outcome has been explicitly confirmed across all types of industries. A rather small number of studies have examined the effect of beauty for occupations, which require good looks. It seems logical that attractiveness plays a crucial role in occupations such as modeling, salespersons, newscasters. For these occupations, perhaps more so than other sectors, the beauty premium should be larger due to interpersonal relationships with employees and customers. In fact, the size of the beauty premium for these occupations is only slightly larger than that estimated for elsewhere (Sachsida *et al.* (2003); Arunachalam & Shah (2012)).

Another important group of studies presents the evidence of the relationship between beauty and productivity for the occupations, which imply face-to-face interaction with employer or customer, but do not require good looks. Biddle & Hamermesh (1995) used the longitudinal sample of graduates of the law school. The authors examined pay differentials in relation to their physical attractiveness. The results have shown that the good-looking lawyers earned more than those who were assessed as less attractive after 5 years of practice, and the positive effect grew with their experience.

Hamermesh & Parker (2003) reported that teaching instructors who are viewed as better looking receive higher instructional ratings by students. The authors have shown that instructional ratings are translated into higher earnings. However, the authors recognized the impossibility of disentangling the productivity-based nature of this outcome from discrimination.

Salter *et al.* (2012) examined the beauty effect on wages of real estate agents and concluded that beauty augments the wages of more attractive agents. The authors emphasized that good-looking agents supplement such characteristics as efforts, organizational skills and intelligence by good looks.

Parrett (2015) conducted a study to investigate the effect of beauty on the earnings of restaurant servers. Using the data on tipping, which was collected outside of Virginia's restaurants, the author found that attractive servers earn approximately 1261 dollars more per year in tips than unattractive servers.

Some research has investigated the relationship between physical attractiveness and academic productivity. There is some evidence that beauty has an impact on the performance of students, the results are consistent with the magnitude of impact founded in the labor market (Ritts *et al.* (1992), Cipriani & Zago (2011), Von Bose (2012), Hernandez-Julian & Peters (2017)). According to Cipriani & Zago (2011), physically attractive students perform better at the exams. The research investigated how students' physical appearance influences their examination results and the authors have concluded that the beauty effect couldn't be attributed to professor's discrimination. Distinguishing between the results of oral and written exams, the authors have provided evidence of a productivity-based explanation of the beauty effect. As opposed to it, Hernandez-Julian & Peters (2017) provided the results of research, which do not support the assumption of higher productivity of more attractive students.

The authors pointed out that students of above-average physical attractiveness get significantly lower grades.

After confirming the existence of the relationship between beauty and productivity for occupations that require face-to-face interaction, the researchers started testing the assumption that attractiveness may enhance a worker's productivity even in jobs, which do not require personal interaction. If beauty is correlated with productivity, it must be supported by the evidence of premiums in a case when the individual cannot be seen.

The influence of the author's attractiveness on research productivity in academic writing was studied by Paphawasit and Fidrmuc (2017). Using the data of 2800 authors who published their works in 16 economic journals the researchers have found the significantly positive effect of an individual's attractiveness on research productivity. The research productivity was measured by citations, journal ranking, and journal impact factors.

Several studies provide some supporting evidence for the relationship between physical attractiveness and labor productivity in individual sport performance. Berri *et al.* (2010) conducted a study using the dataset of National Football League quarterbacks. The authors concluded that physically attractive players are paid greater salaries and this premium persisted after controlling for the player's performance. Ahn & Lee (2013) identified a strong positive effect of beauty on performance-based earnings using the sample of female golfers participating in the Ladies Professional Golf Association tour. Bakkenbuell (2017) found the significantly positive relationship between physical attractiveness and athletic performance irrespective of gender differences. In summary, these findings suggest that sport managers would benefit from hiring more attractive players.

The size of beauty premium depending on gender generate very divisive discussions among researchers. The literature identifies some variation in the effect of attractiveness on labor market outcomes with respect to gender. The results of several studies confirmed that there is no relationship between gender and size of beauty effect on productivity (Fletcher (2009)). Most of the studies, however, report that beauty premium is larger for men than for women. Additionally, the effect is considered independent of customer contact (Hamermesh & Biddle (1993) , Biddle & Hamermesh (1995), Pfeifer (2012)). For instance, attractive men receive twice as high a call-back rate as attractive women (Hamermesh & Biddle (1993)), the good-looking male instructors receive significantly higher ratings from students (Hamermesh & Parker (2003)), more attractive male candidates are more successful in the elections (King & Leigh (2007)).

In contrast to these findings, French (2002) have found the beauty premium for women, though not for men using the data sample for three large organizations in the USA. Doorley & Sierminska (2012) have also indicated the larger beauty premium for women in Europe. There is also a larger beauty premium for female candidates in political performance according to Berggren *et al.* (2010).

If beauty can result in different wage premiums for men and women, it may affect individual labor market responses. Mocan & Tekin (2010) argue that attractive women participate in the labor market since beauty brings them more confidence. Therefore, attractive women are more likely to take advantage of their appearance, if it enhances productivity traits in employment. Avoiding the labor market can be attributed to the fact that less attractive women encounter some barriers to entering the labor market (Hamermesh (2011)). Hence, the smaller beauty premium for women may arise from self-selection into the labor force.

In a series of papers, researchers have examined whether the premium for being attractive and the penalty for being ugly are symmetrical. The literature provides different findings. The number of studies has identified that the effect of beauty is asymmetric, with a greater penalty for unattractiveness

(Hamermesh & Biddle (1993), Hamermesh *et al.* (2002)). Analyzing the US data sample, Hamermesh & Biddle (1993) found that penalties for unattractive workers are greater than premiums for attractive ones. The wage penalty for being unattractive was found to be approximately 11 percent for women and 15 percent for men, which is slightly higher than the beauty premium effect. Using the UK longitudinal data, Harper (2000) found the larger penalty for unattractiveness, which was higher for men. Several studies have reported small evidence of asymmetry in the impact. For example, Hamermesh & Parker (2003) analyzed the Canadian data and found that there is no asymmetry in the beauty effect.

It is evident from the empirical literature that there are no common metrics for beauty or productivity. This fact influences the choice of searching strategy in the context of further meta-analysis. The following sections will provide a brief review of the methods of measuring beauty and productivity in the literature.

2.1 How productivity was measured

Productivity is a widely used economic measure. It identifies a value of output a worker has produced per unit of labor input. Labor and capital are considered as inputs, revenues generally measure output. Labor productivity can be measured in growth rates or levels.

As discussed earlier, the researchers have examined whether the physical attractiveness proxies for unobserved productivity. It means that productivity-enhancing effect may arise when physical attractiveness is a direct determinant of a worker's productivity. Still, in labor economics, there is no direct evidence of the impact of physical attractiveness on productivity. It is not evident that the individual's productivity produces economic benefits as well.

Measuring of productivity of workers can also be used for estimating human capital accumulation. Good-looking workers may have stronger incentives to improve their productivity in the case when prejudice generated by employers and customers provides them with premiums. In such circumstances, the unit produced by an attractive worker is rewarded more than the unit from the unattractive worker. Hence, good-looking workers may put more effort into improving their productivity.

The definition of worker's productivity depends on the data settings. Thus, the productivity of workers can be measured as an output (units or sales produced), relative to an input (number of hours worked or the cost of labor). Labor productivity may be also derived from the aggregate measures at the firm's level as a value-added per worker Pfann *et al.* (2000). Most commonly, studies use the input measures, such as workers' wages, as a measure of productivity at the individual level (Frieze *et al.* (1991), Hamermesh & Biddle (1993), Biddle & Hamermesh (1995)).

There are several reasons why wages do not directly reflect a worker's productivity. Sauermann (2016) notes that such factors as age and tenure can determine the wages, at least partially. These factors, however, depend on institutional settings, which are resulted from collective agreements. Moreover, most of the corporative data do not contain information on hourly wages, but rather on monthly wages. The measures of labor productivity and wages have their weaknesses when it relates to worker's productivity. The most precise measuring method would involve observing productivity for each individual worker at each point in time.

Most occupations, however, have several metrics that can help to evaluate how well workers perform when their occupational duties. Each occupational task might be evaluated along different dimensions, namely, by the quality and quantity of a task. A worker can work quickly, but provide low quality, or slowly, but with high quality.

When evaluating the teacher's performance, for instance, professors are involved not only in teaching but also in research duties and performing administrative tasks. Measuring teachers productivity could be based on the estimation of students' test scores, controlling for student and school characteristics. To determine the teaching productivity, Hamermesh & Parker (2003) have concentrated on the student's assessment of the course.

More direct measures of an individual's productivity are available in academic research. Academic productivity can be measured by the number of publications, but also by the quality of the publications (measured by a journal's impact factor). Sen *et al.* (2010) offers to assess research productivity by the number of publications and facts of co-authorship, citations and grant funding. Paphawasit and Fidrmuc (2017) measured the productivity in academic publishing by the citations, journal ranking, and journal impact factor.

Von Bose (2012), Deryugina & Shurchkov (2015), Talamas *et al.* (2016) measured the student's productivity by taking the grade point average (GPA) across every year of studying, weighted by every module's credit completed by the student. As a determinant of student's productivity Cipriani & Zago (2011) used the integrated index of the number of exams multiplied by the average grade.

Sometimes productivity is only observable at the team level and it is not always possible to estimate the individual contributions to team productivity. For instance, it is the case of team sports, where all performances of individual players strongly depend on the performances of their teammates. Ahn & Lee (2013) investigated the role of beauty for golf players. The authors emphasized that the player's average field score can be a measure of her own productivity. Additionally, the authors used the total prize money earned by a player to identify the effect of beauty on performance-based earnings.

The occupation that is rarely related to productive behavior is a politician. But even for workers from this group, it is possible to construct some performance measures to demonstrate how beauty influences electoral success. One measure of politicians performance is the number (or share) of votes for the candidate (Berggren *et al.* (2010)).

Measuring an individual's productivity is not straightforward, or sometimes even possible. First, there is no universal measure of a worker's productivity. Policymakers and scientists thus use various measures that capture worker's productivity in their specific settings.

2.2 How beauty was measured

It is often stated that beauty is an ascriptive characteristic and it is "on the eye of the beholder". In reality, it is not correct to assume that the definition of beauty is completely subjective. Although beauty is an abstract concept that cannot be measured in absolute terms, observers do agree in specific judgments of who is attractive and who is not.

It must be recognized that beauty standards differ across cultures, but these standards change rather slowly over time. The existence of common standards of beauty was repeatedly confirmed in the empirical literature by finding the substantial agreement among independent raters about the physical attractiveness of individuals (Hamermesh & Biddle (1993), Biddle & Hamermesh (1995), Cipriani & Zago (2011), Ahn & Lee (2013)). The researchers had compared the scores of respondents' beauty in different survey rounds by different observers, and concluded that the raters basically agree with each other. The meta-analysis by Langlois *et al.* (2000) shows that the observers agree on the beauty standards within and between different cultures. Clearly, the observers do not always completely agree on looks, there is often a slight disagreement about individual's attractiveness. To deal with a problem that different raters might have various opinions on physical attractiveness, researchers use standardized beauty ratings.

Facial beauty is the most commonly used measure of physical attractiveness in the literature. Facial beauty seems to be a reliable proxy of physical attractiveness because people form their first impressions from faces: facial beauty is usually associated with a friendly appearance. More than 2000 academic papers on facial beauty were published in the last 30 years.

The authors of several studies argue that beauty is not limited to facial features. The alternative measures of physical attractiveness, which are generally used in economics are height and weight (Frieze *et al.* (1991), Persico *et al.* (2003), Loureiro *et al.* (2012)). The economists argue that height and weight measures are less sensitive to measurement errors, but these measures play a smaller role in the perception of physical attractiveness because beauty cannot be measured only by height and weight standards. The optimal measure of beauty would probably account for all personal characteristics, which are able to form a visual impact on an observer.

Hamermesh & Biddle (1993) used beauty's measurement strategy, which is based on interviewer's ratings of beauty. French (2002) provided the self-reported ratings of beauty. Boo *et al.* (2013) used objective measures of beauty from the literature. The most frequently used approach to assess the physical attractiveness relies on independent photo-based ratings of beauty (Biddle & Hamermesh (1995), Cipriani & Zago (2011), Mobius & Rosenblat (2006), Scholz & Sicinski (2015), Salter *et al.* (2012), Hernandez-Julian & Peters (2017)). The use of photo ratings has a number of specific requirements and settings. First, raters should not be familiar with the assessed person. Second, it is highly important to avoid a relative evaluation instead of evaluating based on the rater's general beauty perception. This issue was mentioned by Biddle & Hamermesh (1995) and the authors suggested to use the photographs copied and mounted on separate sheets of paper. The potential measurement errors may also arise when the raters cannot distinguish beauty from well-dressing or grooming.

The common stylised fact is that beauty ratings are skewed to the right

(Hamermesh & Parker (2003), Hamermesh *et al.* (2002), Ahn & Lee (2013)). The skewness occurs because the raters resist assessing an individual as ”extremely below average”, but they more often give the ”extremely above average” beauty score.

The literature review shows that the findings are different across the studies and there is no guideline value of the effect of beauty on an individual's productivity. The aim of this thesis is to aggregate the core studies revealing the channels through which beauty can influence productivity to clarify the true effect.

Chapter 3

Methodology and Data

3.1 Methodology

3.1.1 Meta-analysis approach

This chapter provides a description of the research methodology used in this thesis. The literature review can provide a summary of the research, and it can help to assess the quality of individual studies on the specified topic. A narrative review, however, is not sufficient when we need to calculate pooled estimates of the effect. The potential problem of narrative review is that the reviewer may include the results which are more preferable from her point of view. The omitted results may cause a distribution bias. The conclusions of the narrative literature review, hence, might be highly affected by the beliefs and expectations of the reviewer. To investigate the true effect it is necessary to review the empirical literature quantitatively.

Meta-analysis introduces a systematic quantitative survey of the empirical literature. This approach is intended to increase statistical power by integrating the findings of empirical research. It focuses on identifying the factors, which could potentially influence the specified effect and hence it is not sensitive to human factors in conclusions.

The first meta-analysis was published in 1904 by the British statistician Karl Pearson in the British Medical Journal. The approach became popular in the medical field: it was an effective tool to aggregate the results of numerous clinical tests. Later, the meta-analysis approach had been introduced in psychology and other fields, including economics. In 1989 Stanley and Jarrell presented a quantitative method of literature survey in economics. Since that date, the average number of published meta-analyses in economics per year had reached the value of 626 in 2012 (Stanley *et al.* (2013)). Over the 30 years of application, the meta-analysis considered valuable for providing objective and complex quantitative reviews of economics research. The most recent meta-analytic research in economics focus on habit formation in consumption (Havranek *et al.* (2017)), economic growth and foreign investments (Gunby *et al.* (2017)).

There are several stages of meta-analytic research according to Stanley (2001). All relevant studies should be considered in the first stage of meta-analysis. Researchers usually use a computer search of standard databases and then reduce the sample of studies to those that contain some relevant estimate and precision measure. The second stage starts with a choice of summary measure (a dependent variable) the researcher wishes to analyze. The dependent variable of meta-regression analysis should be a common comparable metric, transformed from the summary statistic obtained from each study. Researchers usually use regression coefficients, elasticities, t-values, results of statistical tests or effect sizes as a summary statistic. A set of independent variables should be formulated on the third stage of meta-analysis. The set of independent (or moderator) variables includes characteristics that have to be coded into the meta-analysis database: these usually are method characteristics, specification characteristics, data and publication characteristics. The next stage involves the conduct of meta-regression analysis for the purpose of explaining a variation in results among the studies. The meta-regression analysis is intended to answer the question of whether a particular choice of different characteristics influences reported results. If a certain meta-regression variable is found to be significant in a meta-regression analysis, further research in the field should take into account the influence of this variable. It is important for researchers to consider potential sources of excessive variation among the reported estimates. In addition to conventional tests for heteroscedasticity, autocorrelation, and misspecification, researchers should conduct specific tests for publication bias.

To recapitulate, a carefully designed meta-analytic review provides useful information for researchers and policy-makers since it involved a structured and objective examination of the empirical literature. However, there are several limitations to the meta-analytic approach. First, the correctness of the meta-analysis is highly dependent on comprehensive searching and selection strategies. Second, a problem of common misspecification bias may occur: if all included studies contain the same misspecification, the interpretation of meta-analysis cannot be correct.

3.1.2 Publication bias

Publication bias or the ”file drawer problem” is a serious threat to empirical economics. Publication selection exists when researchers prefer results supported by theory or their prior beliefs. Publication bias as a term has been recognized in 1979 in the medical area. Later, the publication selection problem was detected in economics. One of the first studies which identified the presence of publication bias was the research by Long & Lang (1992). Using meta-regression analysis, publication bias has been identified in such areas as minimum wage effects (Doucouliagos & Stanley (2009)), price elasticities (Nelson (2013)), and many others. Nonetheless, there are areas of economic research, where the presence of publication selection has not been detected.

Efendic *et al.* (2011) examined the economic growth literature; the author reported that no evidence of publication bias was presented for the effect of institutions on macroeconomic performance.

Stanley (2005) is associated with a preference for significant results of research. Statistically significant results are perceived as beneficial by researchers. These results are easier to sell and therefore they are more likely to be published in academic journals. Researchers are discouraged to report insignificant results and they often remain in the "file drawer". Because of selective reporting in the empirical literature, significant results are overrepresented, while the true effect can be overestimated or underestimated. Hence, the publication bias brings a gap between the published results and the subject of interest.

The conventional econometric techniques cannot help to address the publication bias. If selection for empirical effect exists for the specific research field, econometric estimates are "overwhelmed": they have skewed or truncated distribution. Taking an average of regression coefficients becomes inefficient for establishing a true effect.

The identification of possible publication bias is the most crucial task of meta-analysis. The meta-analytic approach is found to be robust against publication selection problem and hence it allows to draw precise conclusions. There are several meta-regression methods of testing for a presence of true effect under potential publication selection in the literature. The most important meta-regression tests for publication selection according to Stanley (2007) are funnel-asymmetry testing (FAT), meta-significance testing (MST) and precision-effect testing (PET).

The most commonly used technique to detect publication bias is a graphical analysis, which implies a visual examination of a funnel plot. The funnel plot depicts the inverse value of standard errors (precision) on a vertical axis against the effect sizes on a horizontal axis. Estimates should range randomly around a true effect if studies are not affected by publication selection. The most precise estimates are located very close to the underlying average effect whereas the less precise estimates, which provide larger standard errors, are located at the bottom of the funnel plot. The funnel plot is expected to be inverted and skewed in the presence of publication bias. The asymmetry of the funnel plot may be caused by heterogeneity as well.

Since the interpretation of the funnel plot might be highly subjective, the meta-regression tests have been proposed to measure the relationship between precision and effect size. Description of implementation of meta-regression tests (FAT and PET) can be found in the following chapter as well as the funnel plot of the estimates of beauty effect.

3.1.3 Partial correlation coefficient

Regression coefficients that describe the size and direction of the relationship between physical attractiveness and productivity are of key interest to further analysis. The problem arises from the fact that different studies use different units to measure both variables. Estimates from the selected studies, therefore, are not explicitly comparable. Standardized estimates of the effect size, which allow comparing results of different studies directly, are needed. The modern meta-analyses use partial correlation coefficients to solve this problem (Doucouliagos & Stanley (2009); Efendic *et al.* (2011); Valickova *et al.* (2014)). A partial correlation coefficient is represented by the following equation:

$$pcc_{ij} = \frac{t_{ij}}{\sqrt{t_{ij}^2 + df_{ij}}} \tag{3.1}$$

In the equation 3.1, pcc_{ij} refers to partial correlation coefficient from i^{th} regression's estimate of the j^{th} study; t_{ij} denotes t-statistics of i^{th} regression estimate of the j^{th} study ; df represents corresponding number of degrees of freedom.

To employ the modern meta-analysis techniques, a corresponding standard error for each estimate of the partial correlation coefficient must be calculated. The standard error can be obtained from the previously described estimates, employing the following equation by Fisher (1954):

$$SEpcc_{ij} = \frac{pcc_{ij}}{t_{ij}} \tag{3.2}$$

In the equation 3.2 $SEpcc_{ij}$ is conventional measure of precision, which denotes standard error of the partial correlation coefficient pcc_{ij} ; t_{ij} denotes t-statistics from i^{th} regression estimate of the j^{th} study.

The simple meta-regression model examines the effect of the standard error of the partial correlation coefficient ($SEpcc_{ij}$) on a standardized effect size of the desired effect itself:

$$pcc_{ij} = \beta_0 + \beta_1 SEpcc_{ij} + \epsilon_{ij} \tag{3.3}$$

3.1.4 Heterogeneity

Following the meta-analysis approach, it is necessary to explain the structural heterogeneity that remains after filtering out publication bias. According to the literature review, it is expected, that the effect of beauty may differ systematically across genders, occupations, cultures and geographical regions. These differences might be economically important. Moreover, some variation in estimates can be explained by using different estimation methods.

To investigate why the studies report different estimates of the beauty effect, it is necessary to regress the reported estimates (or their partial correlation coefficients) on explanatory variables, which incorporates possible sources of heterogeneity. This methodology is called meta-regression, which is the most powerful tool of meta-analysis. The meta-regression model can be represented by the following equation:

$$pcc_{ij} = pcc_0 + \beta_j * \sum_j X_{ij} + u_{ij} \tag{3.4}$$

where pcc_{ij} is the partial correlation coefficient of the considered effect estimate; pcc_0 represents the constant; β_j identifies the vector of the coefficients and X_{ij} represents the explanatory variables which capture study characteristics, including the publication bias; u_i is an error term. This specification implies that publication bias, if present, varies randomly across studies and only a systematic variation of the genuine effect is modeled.

After the estimation of the meta-regression model, a procedure of model selection should be performed: the insignificant regressors suppose to be excluded one by one to get a model that contains only significant explanatory variables. With a substantial number of explanatory variables collected from the literature, it might be complicated to choose which variables should be included in the model. Model averaging techniques are the most popular and efficient tools to solve the problem of model uncertainty. The bottom line of these techniques is to regress all possible models with different combinations of variables and then to assign a model's weights. The models, which are better specified, receive the larger weights.

Two averaging techniques are used in this thesis: the Bayesian Model Averaging technique (BMA) and the Frequentist Model Averaging (FMA). BMA technique is the most frequently used in the recent meta-analyses (Eicher *et al.* (2009); Havranek *et al.* (2015); Havranek *et al.* (2017); Havranek *et al.* (2018)). There are several important statistical determinants of the BMA technique. Posterior model probability (PMP) is similar to the information criteria in classical econometrics. Posterior mean (PWM) is an analog of the model average parameter estimate. Posterior standard deviation can be thought of as a standard error. Finally, posterior inclusion probability (PIP) is an analog to statistical significance, which is a sum of PMP for the models in which particular variable is included. The BMA computes posterior means across all models, using the values of posterior model probabilities as weights. While implementing the BMA researchers often apply the Markov chain Monte Carlo algorithm, which uses only the models with high PIP. Computing all the models becomes impossible with a large number of variables. The weight of the prior on individual coefficients (so-called g-prior) must be selected for the BMA procedure as well. The most commonly used g-prior is called the unit information prior (UIP), and it gives the same and very small importance to each coefficient. The most used prior on model probability also gives the same prior weight to each model.

An alternative model averaging technique, which helps to address model uncertainty issue, is called FMA. This technique does not require using prior information, however, the idea to restrict a number of estimated models remains the same as for BMA. To restrict this number, FMA uses Mallow's model averaging estimator and orthogonalization of the covariate space. The weights for model averaging are selected in accordance with the Mallow's criterion. A detailed description of the implementation of these methods can be found in the following chapter.

3.2 Data

According to the approach proposed by Stanley *et al.* (2013) in the "Meta-Analysis of Economics Research Reporting Guidelines", the research had started with searching and collecting of the relevant empirical literature on beauty's effect on productivity. The Collecting of studies went through several stages.

At the first stage, the studies were identified by searching in Google Scholar, RePEc and Scopus databases for any reference to "beauty", and "physical attractiveness", combined with such keywords as "productivity", "performance evaluation" and "discrimination". The abstracts of these works were considered and only those, that contain the empirical estimates, have been collected for further investigation. Additionally, the references and literature reviews of the most-cited studies were checked to expand the list of literature. The overall number of studies at this phase was 76. The search was conducted using English keywords and terminated on March 1, 2019. Only the literature which reports any measure of precision, such as standard errors, t-statistics or p-value was considered for further analysis in order to use the modern meta-analysis techniques and control for the publication bias.

All the studies were revised at the second stage, in order to see whether they include beauty or physical attractiveness rating as an explanatory variable and a proxy of productivity (earnings or performance ratings) as a dependent variable. As it was mentioned in the previous chapter, there are significant differences in measuring both variables.

Studies were considered for inclusion in the meta-analysis only if they deal explicitly with the beauty rating of the respondent. Photo-based evaluation by raters, evaluation of beauty by interviewers and self-evaluated beauty ratings were considered in the analysis. Although the other physical characteristics have been also associated with attractiveness, the studies that are focused on the height and weight of respondents were excluded.

As stated previously, the substantial number of studies use earnings, wages or income as a proxy for productivity measuring. The authors of these studies mostly estimate the relationship between beauty and productivity in the form of the Mincer type earnings function (Harper (2000); French (2002); Fletcher (2009)). In its general form a model is represented by the following equation:

$$ln(Earnings_i) = \beta_0 + \beta_1 Beauty_i + \beta_2 X_i + \beta_3 Y_i + \epsilon_i \tag{3.5}$$

where $ln(Earnings_i)$ denotes the individual level of annual or hourly counted earnings; $Beauty_i$ indicates the individual average attractiveness score; X_i is a vector of individual characteristics such as age, gender, race, country, etc.; Y_i represents the indicator variable, which indicates whether an occupation requires good looks that could enhance productivity; and ϵ_i is the error term. For model 3.4, the positive sign of β_3 can be interpreted as an occupational sorting presence. The occupational sorting hypotheses described by Hamermesh & Biddle (1993) suggests that occupational requirements for attractiveness produce an independent effect on earnings and employees consequently select a certain occupation based on their appearance.

Nevertheless, the dataset was not restricted to the studies that employ the earnings model, since the considered relationship between beauty and productivity can be estimated using different strategies. A substantial number of studies use the occupational-specific performance rating as a determinant of productivity. These studies mainly focus on the analysis of beauty effect for a specific occupation. The researchers use adaptations of the conceptual productivity model formulated by Hershauer & Ruch (1978), which represents productivity as a function of different factors: task capacity, individual capacity, individual effort, and uncontrollable interferences. The model is often represented in the following form:

$$Productivity_i = \beta_0 + \beta_1 Beauty_i + \beta_2 X_i + \beta_2 Z_i + \epsilon_i \tag{3.6}$$

For equation 3.5 $Productivity_i$ denotes the individual productivity in measures of occupation under consideration ; $Beauty_i$ is an individual average beauty score; X_i represents the vector of social determinants such as gender, race, country, marital status, age, etc.; Z_i indicates the vector of occupation-specific characteristics such as team size, tenure, occupational rank, etc; ϵ_i is an error term.

Most of the selected studies report multiple estimates of the relationship between beauty and productivity. According to the recent meta-analytic practices, all estimates given in individual studies were collected, the resulting dataset contains 418 estimates from 37 studies. 32 of these studies are published in the refereed journals, 2 are working papers, 3 papers are the parts of dissertations. The list of studies included in the meta-analysis is presented in Table 3.1.

Table 3.1: List of studies included in meta-analysis

Author(s)	**Year**	**Author(s)**	**Year**
Ahn and Lee	2014	King and Leigh	2009
Anyzova and Mateju	2018	Kraft	2012(a)
Arunachalam and Shah	2012	Kraft	2012(b)
Bakkenbüll and Kiefer	2015	Leigh and Borland	2007
Bakkenbüll	2016	Oghazi	2016
Bakkenbüll	2017	Oreffice and Quintana-Domeque	2016
Berri et al.	2011	Paphawasit and Fidrmuc	2017
Biddle and Hamermesh	2008	Pfann	2000
Borland and Leigh	2015	Pfiefer	2012
Cipriani and Zago	2011	Ponzo and Scoppa	2012
Dossinger et al.	2019	Sachsida et al.	1994
Fletcher	2009	Salter et al.	2012
French et al.	2009	Scholz and Sicinski	2015
Hamermeshand Biddle	1993	Sen et al.	2016
Hamermesh and Parker	2005	Tao	2008
Hamermesh et al.	2002	Walcutt	2011
Harper	2000	Wolbring and Riordan	2016
Julian and Peters	2017		
Kanazawa and Still	2017		

The earliest study included in the meta-analysis has been published in 1993 (Hamermesh and Biddle), and the latest one has been published in 2019 (Dossinger *et al.* (2019)). Based on the Google Scholar citation numbers, the most cited papers are Hamermesh & Biddle (1993) - 1653, Biddle & Hamermesh (1995) -510, Hamermesh & Parker (2003) -440 and Harper (2000) -375. The authors include approximately a similar set of control variables in their estimations, which mainly consists of variables of individual characteristics of respondents. Most of the studies control for age of respondent (14 from 37), experience (18 from 37) and education (15 from 37). Several studies control for race, gender and health status of respondents as well. The highest number of studies examines beauty bias for the Europeans, Americans, and Canadians. There are several studies that use datasets with mixed nationalities of respondents.

Of the 418 estimates of the beauty effect on productivity in the sample, 186 coefficient estimates are positive and statistically significant,129 are positive but insignificant, 10 are negative and significant, and 89 are negative but insignificant. As it can be observed, the authors exploit different measures and scales in their analysis. Since the estimates are heterogeneous and hardly comparable, the following statistics and estimates are obtained using partial correlation coefficients from each effect size estimate.

The mean reported estimate of the beauty effect is 0,073, the mean value weighted by the inverse number of estimates per study is 0,097. However, the beauty effect size is much smaller against the results of the meta-analysis of experimental studies (Hosoda *et al.* (2003)) even before the checking for publication selection. Doucouliagos (2011) provided the guidelines under which the partial correlation coefficient in the range between 0.07 and 0.17 in absolute value is considered "small", hence the partial correlation coefficient of 0,073 represents a small effect of beauty on productivity.

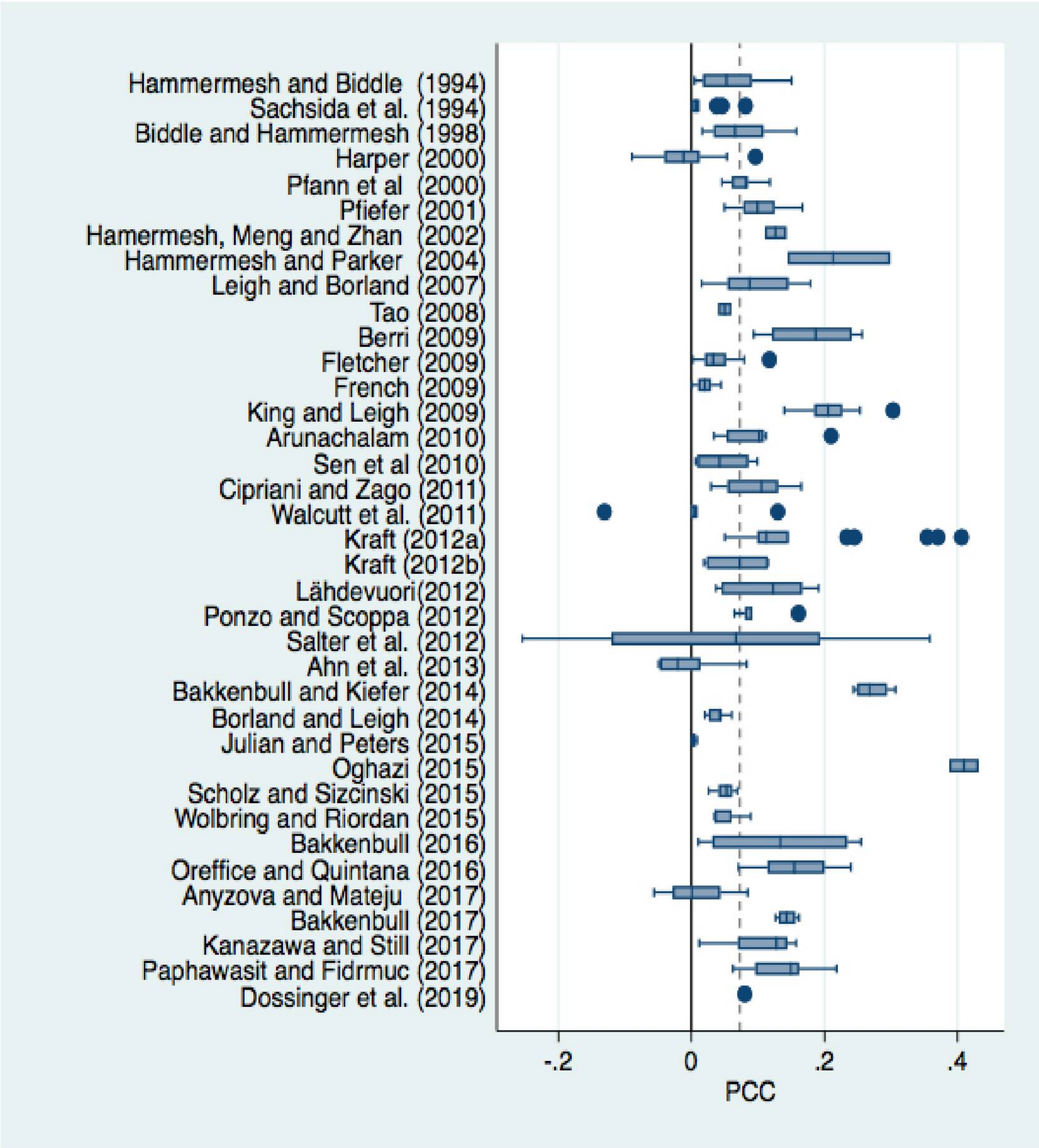
Table 3.3 presents the summary statistics for the partial correlation coefficients for different subsets of data. The dataset demonstrates substantial heterogeneity in terms of the methodology employed, geographical region and time period covered. Beauty also differs by gender. To display the findings of studies included in the meta-analysis graphically, the forest plot technique is widely used. The forest plot in Figure 3.1 shows that the partial correlation coefficients of estimates of beauty effect are not homogeneous and differ across and within studies.

Table 3.2: PCC of beauty effects for different subsets of data

	N	Mean	St.dev	Min	Max
Productivity type					
Earnings-based	283	0.067	0.078	-0.089	0.406
Performance-based	135	0.084	0.101	-0.254	0.432
Gender of respondents					
Males only	151	0.057	0.070	-0.089	0.299
Females only	149	0.056	0.074	-0.130	0.307
Both genders	118	0.114	0.105	-0.254	0.432
Beauty assessment					
Self-rated	19	0.034	0.066	-0.13	0.179
Interviewer-rated	203	0.047	0.065	-0.089	0.257
Multiple raters	196	0.103	0.097	-0.254	0.432
Geographical region					
Europe	123	0.057	0.087	-0.089	0.432
North-America	169	0.055	0.068	-0.254	0.358
Others	126	0.112	0.096	-0.130	0.406
Mixed nationalities	44	0.115	0.096	-0.049	0.307
Occupation type					
Dressy occupations	138	0.082	0.096	-0.25	0.432
Other occupations	82	0.085	0.088	-0.049	0.307
Estimation type					
OLS	329	0.069	0.089	-0.25	0.432
Other estimators	89	0.087	0.075	-0.049	0.307
Decades of publication					
1990	77	0.057	0.044	0.005	0.158
2000	155	0.054	0.081	-0.089	0.303
2010	186	0.095	0.098	-0.254	0.432
Publication status					
Published studies	323	0.064	0.085	-0.25	0.432
Unpublished studies	95	0.101	0.085	-0.130	0.406
All estimates	418	0.073	0.087	-0.254	0.432

Notes: The table reports mean values of the partial correlation coefficients for different subsets of data. OLS = ordinary least squares. St.dev= Standard Deviation

Figure 3.1: Forest plot



Notes: The figure shows a forest plot of the estimates of beauty effect reported in empirical literature. The boxes on the graph represent the interquartile range ($P_{25} - P_{75}$), the median is marked. Whiskers show the interval from ($P_{25} - 1.5 * interquartilerange$) to ($P_{75} + 1.5 * interquartilerange$) if such estimates exist. Dots show the outliers reported in each study.

Chapter 4

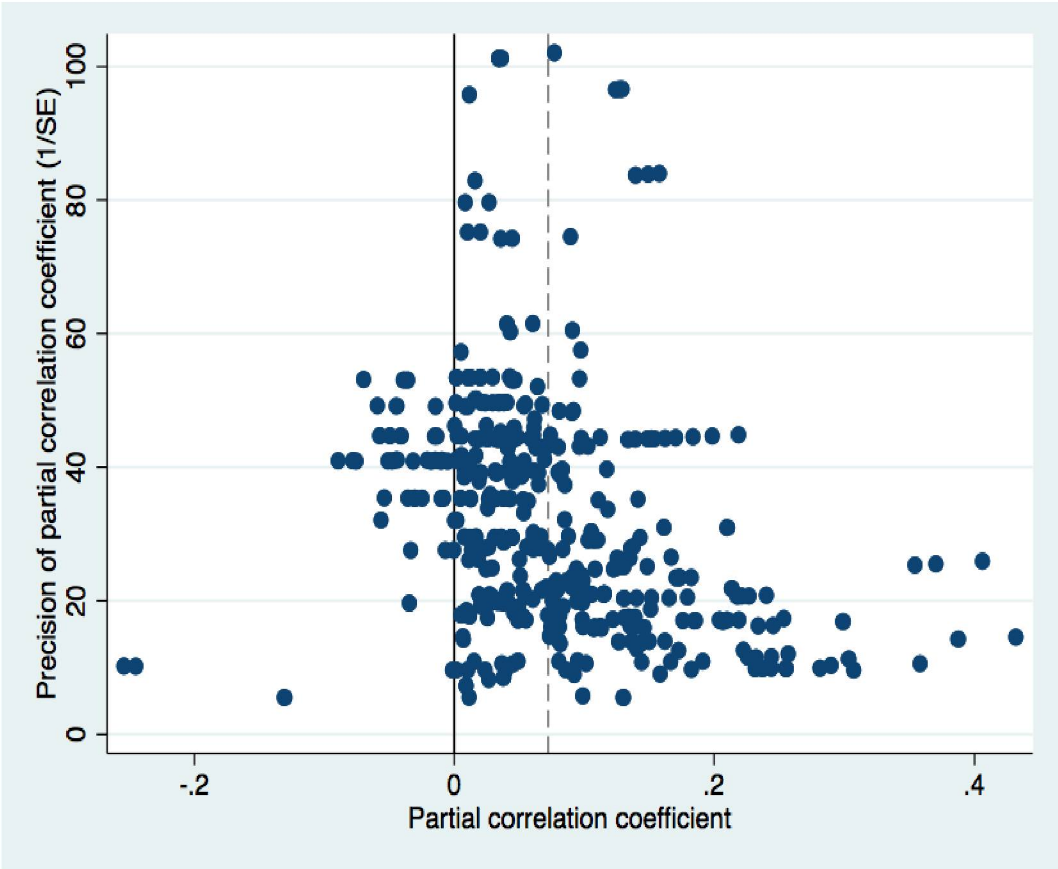
Empirical Results

4.1 Publication Bias

The analysis starts by investigating the presence of publication selection using a funnel plot. The funnel plot depicts partial correlation coefficients, derived from individual estimates of beauty effect on the horizontal axis and the inverted standard errors (as measures of precision) on the vertical axis. This graphical technique allows observing the presence of publication bias, which occurs when researchers prefer a certain direction of results. In the absence of publication bias, the plot should look like a symmetrical inverted funnel: the highly precise observations will be concentrated close to the true effect, while the less precise observations should be more dispersed.

The funnel plot for 37 studies is depicted in Figure 4.1. In order to improve the representativeness of results the estimates with extreme precision values (higher than 120) were excluded from the plot, all estimates will be included in the meta-regressions onward. The funnel plot suggests a small positive true effect, but it does not resemble a funnel and shows an imbalance in the reported beauty effects, as a right-hand tail of the funnel appears to be heavier. There is a significantly lower number of estimates on the left half of the plot. Hence, the positive estimates are preferably selected for publication. This finding supports a prevailing theoretical view of the positive relationship between beauty and productivity.

Figure 4.1: Funnel plot



Notes: The horizontal axis shows the beauty effect on productivity represented by the partial correlation coefficients. The vertical axis shows the precision represented by the inverted value of standard errors. The dashed vertical line demonstrates a zero partial correlation coefficient of the beauty effect on productivity; the solid vertical line demonstrates the mean partial correlation coefficient

Figure B.1, which separates the earnings-based estimates and performance-based estimates, can be found in the Appendix section. The earnings-based estimates are mostly scattered around zero, while the performance-based estimates are scattered to the right from zero. Both plots seem asymmetrical with the heavier right-hand tails. Visualization of the estimates helped to obtain a general picture of the bias, which might be present in the literature. However, the visual method is quite subjective when testing for the publication bias and underlying value of the beauty effect. Therefore, formal testing methods have been applied. Tables 4.1-4.2 summarize results of the following regression:

$$PCC_{ij} = PCC_0 + \beta_1 * SEPCC_{ij} + \mu_{ij} \tag{4.1}$$

where PCC_{ij} denotes i_{th} partial correlation coefficient of the beauty effect estimated in the j_{th} study and $SEPCC_{ij}$ denotes the corresponding standard error. The PCC_0 represents the underlying genuine effect absent publication selection bias. The coefficient of standard error (β_1) identifies the direction and magnitude of the publication bias; μ_{ij} is an error term. If the null hypothesis of $\beta_1 = 0$ is rejected, there is the evidence for funnel asymmetry. The direction of bias is determined by a sign of the β_1 estimate. A statistically significant estimate of the intercept PCC_0 indicates that, on average, there is the true effect of beauty on productivity.

Four specifications, which allow mitigating the problem of potential heteroscedasticity of the error term, have been applied for testing. The first column of Table 4.1 reports the ordinary least squares estimates clustered at the study level. The second column reports the estimates, which use the inverse value of the number of observations as an instrument for the standard errors. The third column shows the estimates of the fixed effect model with the additional random term. The fourth column reports the estimates of the between effects model, which uses only the median estimates of the beauty effect.

The estimates in Table 4.2 were weighted by the inverse value of the standard errors (the columns (1) and (2)). Using these precision weights has enabled to assign greater importance to more precise estimates and to correct the models for heteroscedasticity, as well. The estimates were also weighted by the inverse number of observations per study (columns (3) and (4) of Table 4.2) in order to treat the large and small researches equally.

According to the Funnel Asymmetry Tests(FAT), the publication bias is statistically significant for the unweighted estimates of the beauty effect (Table 4.1). Three of the four estimated models indicated the presence of positive publication bias. Precision Effect Tests (PET) indicated the significant underlying effect of beauty for three of the four models as well. The FAT results for the weighted sample show that models weighted by the inverse value of the standard errors (WLS and IV) indicated the presence of strong positive publication bias for the estimates of the beauty effect. The coefficients of true effect are statistically significant for the OLS and IV models, weighted by the number of estimates.

We also employed alternative methods of correcting for the publication bias in order to check the robustness of our previous results. First, the "Top10" method, introduced by Stanley *et al.* (2010), was used. This method suggests that using 10 percent of most precise estimates for calculations gives more efficient results than summary statistics. The average beauty effect of the most precise estimates is 0.035, which implies that the magnitude of the publication bias is commensurate with previous results of the meta-regression tests. Second, a recent non-parametric stem-based method by Furukawa (2019) was employed to correct the publication bias. The method generalizes the "Top10" technique and relates to the stem of the funnel plot. The result for the beauty effect estimates is 0.02, which is even lower than the results of the "Top10" estimation.

The applied meta-regression tests suggest that all heterogeneity in the results is due to the publication bias and statistical sampling error. However, it is not realistic, and we need to check whether the heterogeneity of the beauty effect is attributable to the use of different study design characteristics.

Table 4.1: Tests of publication bias and true effect

	(1) OLS	(2) IV	(3) FE	(4) BE
PCCSE (publication bias)	0.820**	0.886*	-0.123	0.684*
	(0.46)	(0.48)	(0.23)	(0.45)
Constant (effect absent bias)	0.041*	0.038	0.082**	0.065**
	(0.02)	(0.03)	(0.01)	(0.03)
N (number of estimates)	399	399	399	399

Notes: The table reports the results of testing for the publication bias. The estimates with precision >120 excluded. OLS = ordinary least squares. IV = the inverse value of the number of observations is used as an instrument for the standard error. FE = study-level fixed effects. BE = between effects. Standard errors in parentheses are robust and clustered at the study level. * P < 0.10, ** P < 0.05, *** P < 0.01

Table 4.2: Tests of publication bias and true effect (weighted sample)

	Precision		Study	
	(1) WLS	(2) IV	(3) OLS	(4) IV
PCCSE (publication bias)	1.172**	1.163**	0.541	0.665
	(0.45)	(0 .49)	(0.46)	(0.49)
Constant (effect absent bias)	0.026	0.026	0.073***	0.067***
	(0.02)	(0.02)	(0.02)	(0.02)
N (number of estimates)	399	399	399	399

Notes: The table reports the results of estimates, weighted by the precision or study. The estimates with precision >120 excluded. Study = the model is weighted by the inverse of the number of estimates per study. Precision = the model is weighted by the inverse of the standard error of an estimate. WLS = weighted least squares. Standard errors in parentheses are robust and clustered at the study level. * P < 0.10, ** P < 0.05, *** P < 0.01.

4.2 What Explains Differences in Beauty Bias Estimates

Estimates of the beauty effect vary substantially across the studies as demonstrated in the previous chapters: heterogeneity of the estimates is obvious from the forest plot (Figure 3.1) and summary statistics (Table 3.2). To investigate systematic patterns in the heterogeneity of the beauty effect, the multivariate meta-regression should be employed. The meta-regression equation 4.1 was augmented by a vector of the collected variables which potentially influence the reported beauty effect estimates.

$$PCC_{ij} = PCC_0 + \beta_j * \sum_j X_{ij} + u_{ij} \tag{4.2}$$

where PCC_i is the partial correlation coefficient of the beauty effect estimate; PCC_0 represents the constant; β_j identifies a vector of the coefficients and X_{ij} represents the set of explanatory variables which capture data, estimation and publication characteristics, including the standard error (publication bias); u_i is an error term. The publication bias, if present, varies randomly across studies and only a systematic variation of the true effect modeled.

According to the literature reviewed, the effects of beauty on productivity differ across genders, occupations, cultures and geographical regions. Considering the fact that some of these differences may determine the magnitude of the beauty effect and hence, may produce heterogeneity of the reported results, we have collected 21 characteristics reflecting the data, methods, specifications and publication status from each primary study.

The explanatory variables, that are collected to explain heterogeneity were grouped into the following blocks: 1) Data specifications and characteristics, 2) Variable definitions, 3) Estimation characteristics, 4) Publication characteristics. The following subsections comprise a description of the factors that can contribute to explaining the heterogeneity among the estimates. Table 4.3 provides a description of the 21 explanatory variables with their simple means, standard deviations, and means weighted by the inverse of the number of estimates reported in individual studies.

Data Specification and Characteristics

Most of the studies under review rely on independently pooled cross-sectional data. Several studies, however, have used longitudinal data to examine the relationship between beauty and productivity over time. Hence, a dummy variable for the studies that rely on panel data (*Panel*) included in the list of explanatory variables; the reference category represents the studies, which used the cross-sectional data.

Cultural differences in the evaluation of beauty may cause some variation of the beauty effect across geographical regions and countries. Thus, it is necessary to examine whether geography induces a systematic difference in the estimated beauty effect. Beauty ratings of the respondents from fifteen countries examined in the literature. Approximately one-third of all reported estimates obtained for the US respondents; another third obtained for the European respondents. Therefore, the dataset divided by 5 regions, and the regional dummy variables *Europe* and *North America* included in the analysis instead of the underlying characteristics of the countries.

Differences in the magnitude of beauty effect for males and females have been widely discussed in the literature. The dummy variables *Male* and *Female* introduced to the meta-regression model. The reference category represents the studies, which use a combined group of respondents. Hosoda *et al.* (2003) examined the relevance of gender differentials in the beauty effect in the meta-analysis of experimental studies. The authors consider that the variation of beauty effect across genders explained by the "lack of fit" theory introduced by Heilman (1983).

The same theory predicts that physical attractiveness would interact with a "dressy" occupation. When controlling for the "dressy" type of occupation, we proceeded from the assumption that beauty might be more important for the jobs with more frequent face-to-face interactions. Hence, the list of explanatory variables includes the dummy variable *Dressy*. The occupations divided into dressy and non-dressy categories based on the set of occupations presented by Hamermesh & Biddle (1993).

Researchers in the field of labor economics often control the models for such individual characteristics as age, education and job experience of the respondents. Following their experience, we include dummies *Age*, *Experience* and *Education* to the meta-regression model. Education and experience suppose to have a positive effect on an individual's productivity: more educated and experienced employees should be more productive. The effect of controlling for age is not straightforward. On the one side, an employee becomes more experienced with age and hence more productive. On the other side, some physical possibilities become lower with age, that might be an important factor for some occupations.

The increasing number of research shows that other individual characteristics such as communication skills, IQ tests, leadership skills, confidence, and grooming can correlate with beauty scores and hence can enhance labor productivity. Controlling for these characteristics have confirmed their importance in the most cases (Langlois *et al.* (2000), French (2002), Fletcher (2009)). Therefore, the dummy variable *Cognitive Skills* is included in the list of potentially influencing factors for the beauty effect.

Variables Definition

Despite we transformed the estimates into the partial correlation coefficients, some systematic deviations might remain untreated because of using different productivity and beauty measuring approaches. As already discussed in the previous sections, there are two common ways to assess occupational productivity in the literature. Mean reported estimates of the beauty effect obtained by using the earning-based model differ from the estimates obtained by using the performance-based model (Table 3.2). The dummy variable *Performance-based* is introduced to determine whether the difference holds after controlling for other aspects of data. The reference category represents earnings-based estimations.

Reported estimates may also differ depending on the type of beauty's rating used for research. A large number of studies use standardized beauty ratings obtained from multi-raters evaluations. The other studies use beauty ratings obtained from the self-evaluations of the respondents. The ratings obtained from multi-raters evaluations tend to skew to the right, while the ratings obtained from self-evaluations of the respondents are generally lower. The categories of beauty measuring transformed into the *Multiple Raters* and *Self-Evaluation* dummies. The reference category represents the studies, which used beauty ratings obtained from one interviewer.

Controlling for the data homogeneity might be an important issue for the estimations. To identify potential sources of the beauty effect, a substantial number of studies use the data of employees from a relatively homogeneous group (the same occupation). The dummy *Homogenous*, therefore, was included in the list of controls for the meta-regression model. The inclusion of the *Log* dummy aims to control for log transformations of the dependent variable.

Estimation Characteristics

Researchers use various techniques to estimate the relationship between beauty and individual productivity. Most studies estimate the beauty effect by using linear regression and OLS, although some of the studies assume heteroscedasticity and employ TSLS (Kraft, P. (2012)) and the quantile regression (Paphawasit and Fidrmuc (2017)). A few panel studies use the random effects model for estimation (Ahn & Lee (2013)). Overall, the eight estimation techniques may potentially drive differences in results. However, most of it have been used only once, for particular research. The dummy variable *OLS* introduced to the meta-regression model.

Publication Characteristics

To account for the methodological innovations, the number of modern meta-analyses (Valickova *et al.* (2014); Havranek *et al.* (2018); Havranek *et al.* (2017)) include the year of publication in meta-regression model. The reason can be explained by the fact that advanced methodological and estimation techniques are more likely to cover the unobserved data characteristics, which can affect the reported results. Hence, the *Publication Year* is included in the list of explanatory variables to control whether the role of the beauty impact on productivity has changed over time. To consider the quality of research, we use another two publication characteristics as explanatory variables in the meta-analysis. The variable which counts the number of citations in Google Scholar (*Citations*) is introduced to assess how often the research used as a reference in the literature. The variable *Published* indicates that the study published in academic journals.

4.2.1 Bayesian Model Averaging

With a relatively high number of explanatory variables collected from the empirical literature, the effective methodological tool is needed to analyze the sources of heterogeneity. With 21 explanatory variables, 2^{21} different models could be estimated, but we need to determine the most relevant set of explanatory variables to avoid redundancy. Following the most recent meta-analyses (Havranek *et al.* (2015); Havranek *et al.* (2017); Havranek *et al.* (2018)) the BMA technique is implemented in this thesis.

Before applying the BMA, all estimates have been weighted by the number of observations per study. Since the results of previous research have shown the well predictive performance of the combination of UIP and the unit information

Table 4.3: Description and summary statistics of explanatory variables

Variable	*Description*	*Mean*	*SD*	*WM*
Beauty PCC	Partial correlation coefficient derived from the estimate of beauty effect	0.073	0.086	0.097
Standard Error	The estimated standard error of the beauty effect estimate	0.041	0.029	0.05
Data Characteristics				
Panel	=1 if panel dataset is used	0.07	0.255	0.125
Male	=1 if the estimates of the study are for male respondents only	0.361	0.481	0.273
Female	=1 if the estimates of the study are for female respondents only	0.356	0.480	0.339
Age	=1 if the estimation controls for age of the respondent	0.282	0.451	0.323
Experience	=1 if the estimation controls for job experience of the respondent	0.567	0.496	0.455
Education	=1 if the estimation controls for education of the respondent	0.447	0.498	0.364
Cognitive	=1 if the estimation controls for cognitive skills of the respondent	0.447	0.498	0.364
Dressy	=1 if the concerned occupation requires good looks or or based on social interactions	0.330	0.471	0.405
North America	=1 if the beauty effect estimated for US/Canada	0.404	0.491	0.324
Europe	=1 if the beauty effect estimated for EU countries	0.294	0.456	0.297

Notes: SD = standard deviation, SE = standard error, WM = mean value weighted by the inverse of the number of estimates per study

Table 4.4: Description and summary statistics of explanatory variables

Variable	*Description*	*Mean*	*SD*	*WM*
Variables Design				
Performance-based	=1 if the dependent variable is performance-based	0.323	0.468	0.419
Log	=1 if logarithmic transformation is applied in model	0.722	0.448	0.670
Homogenous	=1 if the study use homogenous group of respondentst	0.871	0.336	0.801
Raters Evaluation	=1 if the beauty is assessed by group of raters	0.469	0.500	0.568
Self-Evaluation	=1 if the beauty is assessed by respondent	0.469	0.500	0.568
Estimation Characteristics				
OLS	=1 if OLS estimator is used to examine the beauty effect	0.871	0.336	0.801
Publication Characteristics				
Publication Year	Logarithm of the publication year	7.604	0.004	7.605
Citations	Logarithm of the number of Google Scholar citations (on Dec,2018)	3.566	2.336	2.659
Published	= 1 if the study is published in a journal	0.773	0.420	0.730

Notes: SD = standard deviation, SE = standard error, WM = mean value weighted by the inverse of the number of estimates per study g-prior, we used this combination as well. The BMA procedure was performed using the BMS package in the R software environment. The results of the BMA estimation are visualized in Figure 4.2. The numerical representation of the BMA results are represented in the left-hand panel of Table 4.5.

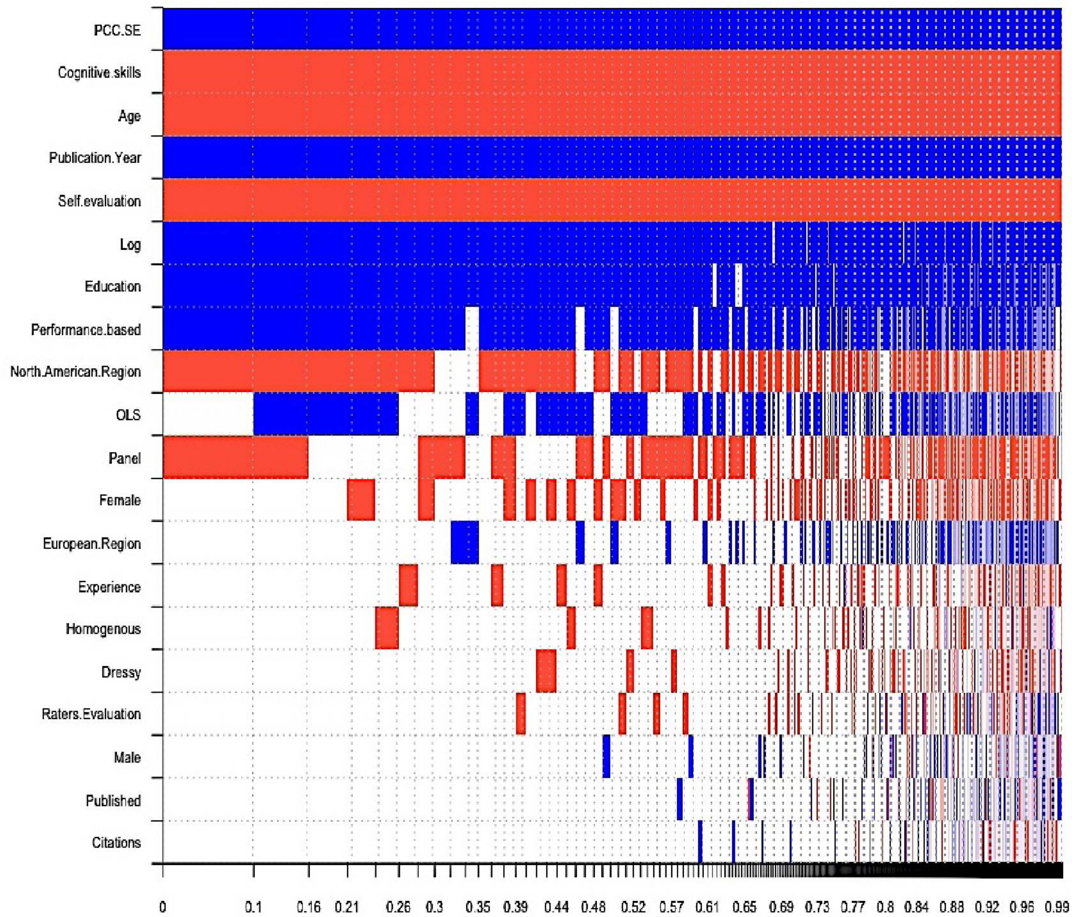
The columns in Figure 4.2 represent the processed models, which are arranged from left to right in descending order. The models are sorted according to their inclusion probability. The rows display explanatory variables, which are arranged from top to bottom in descending order. The variables have been sorted according to their posterior inclusion probability (PIP). In this way, each cell in Figure 4.2 displays a specific variable in a specific model. Each blue-colored cell shows that the variable was included in the model and that the sign of the estimated coefficient is positive. Each red-colored cell indicates that the variable was included in the model and that the sign is negative, respectively. The blank cells reveal that the variables were not included in the model.

The estimation report of BMA includes the values of three underlying statistical measures. First, the Posterior Inclusion Probability (PIP) shows the posterior probability of inclusion of a particular variable in a model. A higher value of PIP is attributed to the higher importance of particular variables when explaining the heterogeneity. Second, the Weighted Posterior Mean (WPM) represents an analog of the model average parameter estimate. The third measure is the Weighted Posterior Variance, which represents the analog of standard deviation.

The principles of interpretation of posterior inclusion probability were formulated by Jeffreys (1961). Jeffreys considers the PIP values between 0.5 and 0.75 as weak, values between 0.75 and 0.95 as positive, values between 0.95 and 0.99 as strong, and values above 0.99 as decisive evidence for an effect. Hence, the results represented in Table 4.2 show decisive evidence of an effect in the cases of *Standard Error*, *Publication Year*, *Self Evaluation*, *Age* and *Cognitive Skills*; strong evidence of an effect in the cases of *Log* and *Education* variables;

positive evidence of an effect for *Performance Based Productivity* variable; and weak evidence of an effect in the cases of the *OLS*, *Panel*, *North American Region* variables.

Figure 4.2: BMA visualisation



Notes: The figure represents the results of the BMA. The vertical axis depicts the explanatory variables ranked according to their PIP in descending order. The horizontal axis depicts the values of the cumulative posterior model probability. The blue color of the cells shows the estimated parameter of a relative variable is positive. The red color of the cells shows the estimated parameter of a relative variable is negative. A cell without color shows the related variable not included in the model

The following approach to assess the remaining heterogeneity is based on a frequentist check. The frequentist check includes the explanatory variables from BMA with posterior inclusion probability higher than 0.5. This specification estimated by OLS with robust standard errors clustered at the study level. The results of the frequentist check estimation can be found in the right-hand panel of Table 4.5. The results show that all explanatory variables from the BMA, except the *OLS* and *Panel*, are statistically significant at 5 percent level.

4.2.2 Robustness Check and Results Discussion

Havranek *et al.* (2017) first applied an alternative practice of model averaging in the meta-analysis. The methodology is called FMA, and it is considered as a robustness check for the results of BMA and OLS models previously applied. The FMA specification includes all collected explanatory variables. Before applying the FMA, all estimates were weighted by the inverse of the number of estimates per study. The FMA procedure was performed in the R software environment. The results of the FMA can be found in Table 4.6, and it shows that the results are predominantly in line with the BMA exercise except for the case of *OLS* explanatory variable: the evidence for the effect of OLS inclusion is not significant according to the estimations.

Resulting from all specifications and estimation methodologies, the evidence for the publication bias remains after the inclusion of explanatory variables. The coefficient on the *Standard Error* is robustly significant when we control for 20 additional factors related to studies and estimates. This finding corroborates our previous results.

Results for Design of Variables

The evidence for the positive effect of inclusion of the variable *Performance-based productivity* is significant for both model averaging approaches and the frequentist check. This result suggests that the choice of proxy for productivity measuring is relevant for beauty effect investigation. Researchers who prefer to use performance-based measures of productivity, and therefore more homogeneous data from specific occupations, obtain higher estimates.

The definition and design of our independent variable (*Beauty*) is also rel-

Table 4.5: Explaining heterogeneity in the estimates of beauty and productivity relationship

Response Variable:		*BMA*			*FC*	
est of beauty effect	*PIP*	*PM*	*PSD*	*Coef*	*SD*	*p-val*
Constant	1	-59.944	NA	-53.893	16.586	0.001
Standard Error	1	0.935	0.141	0.897	0.240	0.001
Variables Design						
Performance-based	0.875	0.234	0.122	0.246	0.077	0.003
Raters Evaluation	0.103	-0.007	0.049			
Self Evaluation	0.999	-0.559	0.120	-0.590	0.145	0,000
Log	0.983	0.345	0,106	0.341	0.09	0.001
Data Characteristics						
Dressy	0.108	-0.012	0.053			
Panel	0.530	-0.128	0.140	-0.22	0.145	0.132
Age	0.999	-0.453	0.083	-0.474	0.141	0.002
Experience	0.145	-0.018	0.056			
Education	0.964	0.289	0.102	0.276	0.112	0.019
Male	0.069	0.003	0.027			
Female	0.281	-0.039	0.073			
North America	0.989	-0.249	0.066	-0.054	0.004	0.000
Europe	0.193	0.028	0.068			
Estimation Characteristics						
OLS	0.555	0.101	0.106	0.146	0.086	0.099
Publication Characteristics						
Publication Year	0.999	7.886	1.086	7.6747	2.180	0.001
Citations	0.0523	0	0			
Published	0.061	0.001	0.021			

Notes: BMA= Bayesian Model Averaging, FC= Frequentist Check, PIP= Posterior Inclusion Probabilty, PM= Posterior Mean, PSD= Posterior Standard Deviation, Coef= OLS coefficient, SD= Standard Deviation, *p*-val= P-value. In the frequentist check we include only variables with PIP higher than 0.5

Table 4.6: Explaining heterogeneity in the estimates of beauty and productivity relationship. Frequentest Model Averaging

Response Variable: estimates of beauty effect	*Coefficient*	*StDev*	*p-value*
Constant	-60.284	16.974	0
Standard Error	1.006	0.151	0
Variables Design			
Performance-based	0.449	0.135	0.001
Raters Evaluation	-0.136	0.122	0.263
Self Evaluation	-0.587	0.125	0
Log	0.314	0.093	0.001
Data Characteristics			
Dressy	-0.112	0.119	0.347
Panel	-0.141	0.107	0.188
Age	-0.462	0.085	0
Experience	0.109	0.158	0.488
Education	0.312	0.105	0.003
Male	-0.021	0.099	0.832
Female	-0.151	0.089	0.093
North America	-0.220	0.093	0.018
Europe	0.03	0.095	0.749
Estimation Characteristics			
OLS	0.218	0.092	0.019
Publication Characteristics			
Publication Year	7.931	2.241	0
Citations	0	0.004	0.978
Published	0.077	0.093	0.408

Notes: The table shows the results of the FMA. Mallow's criterion is used to select the optimal weights for modeling. The number of models reduced using orthogonalization of the covariate space.

evant in determining the sources of heterogeneity of beauty effect estimates. All estimated specifications confirm the importance of inclusion of the *Self-Evaluation* explanatory variable. The sign of the coefficient is negative: the respondents usually understate self beauty ratings. Hence, the studies which use self-evaluated beauty ratings tend to report significantly lower estimates of beauty effect.

Results for data characteristics

According to the results, a coefficient of *Dressy* explanatory variable is insignificant. This finding implies that a type of occupation does not systematically affect the reported beauty effect estimates. Hence, the estimates of beauty effect for occupations, which require good looking produce commensurate beauty effect in comparison with other occupations. This conclusion supports the findings of Kraft, P. (2012), Arunachalam & Shah (2012), Paphawasit and Fidrmuc (2017).

The use of panel data for estimations does not prove important in the FMA and frequentist check specifications. The BMA reported weak evidence for the panel data effect on beauty estimates. The low availability of longitudinal data on physical attractiveness supplemented with economic characteristics might be a reason for a considerably smaller number of studies that use panel datasets. However, it seems to be important to study a beauty effect over time.

Other important factors that produce the heterogeneity of reported estimates are the individual characteristics of respondents, namely, *Age*, *Education* and *Cognitive skills*. The strong negative effect on the magnitude of the beauty effect is attributed to the *Age*. In contrast, the inclusion of the *Education* variable leads to the increase of the beauty effect. Controlling for *Cognitive skills* substantially reduces the magnitude of the beauty effect. This finding is in line with previous results of Mobius & Rosenblat (2006), Salter *et al.* (2012), Scholz & Sicinski (2015). The authors concluded that the magnitude of the beauty effect is decreasing after the inclusion of cognitive characteristics such as IQ tests, communication skills, measures for confidence and personality; however, the beauty effect does not vanish.

According to the results of the BMA and FMA exercises, there is no evidence of significant differences in the beauty effect estimates attributed to gender. This finding suggests that it does not matter whether the authors use male sample, female sample or mixed samples of respondents. This finding is not in line with the previous results: the relationship between physical attractiveness and labor outcomes has been shown to be different for men and women by Biddle & Hamermesh (1995), French (2002) and others. However, the beauty premium gap across genders is expected to decrease due to the raised participation of women in the labor market.

The estimation results regarding the regional differences in the beauty effect are mixed. The estimates of the beauty effect for respondents from European countries do not differ significantly, while the estimates for respondents from the North American region seem to be lower than those for other countries. This finding suggests that the respondents from the US and Canada experience a smaller beauty effect. The possible explanation is that the US and Canada have modern economies, where social orientation plays an important role. Information on the series of protection measures for employees in the US supports this statement (the city of San Francisco in 2001 and the District of Columbia in 2008).

Results for estimation characteristics

Our analysis suggests that researchers who prefer to use OLS estimator, obtain higher values of beauty effect in comparison with the authors who use other estimation techniques. However, the evidence on the importance of OLS using is weak and non-consistent across different model averaging approaches. It seems logical that more advanced estimation techniques would provide more accurate estimates of the beauty effect. After decades of studying, there is no agreement on the best estimation technique for the beauty effect among researchers.

Results for publication characteristics

The additional results related to publication characteristics are important. The first result is the high posterior inclusion probability of the variable *Publication Year* in the BMA and FMA models. A time period when the study was published matters for the magnitude of the beauty effect. According to the results, the coefficient of the variable *Publication Year* is significant and positive. It means that the most recent studies report systematically higher results, which is consistent with the values of partial correlation coefficients estimated for various decades (Table 3.2). The use of publication year may reflect the changes in the estimation approaches and methodologies applied. However, this finding does not meet our expectations and require further study in the longer term. The aspects of research quality that are captured by the other two proxies (*Publication Status* and *Citations*) do not systematically affect the estimates of the beauty effect.

Chapter 5

Concluding remarks

This thesis conducted a quantitative synthesis of 418 estimates of the effect of beauty on productivity as reported in 37 studies. This is the first meta-analysis on the relation between beauty and productivity to the best of our knowledge.

In order to avoid misleading interpretations of the beauty effect on productivity from potentially biased results from empirical literature, we carefully tested the beauty effect for publication selection. We used informal testing of the funnel plot as well as formal testing methods. The results suggested that the estimates of beauty effect are influenced by publication bias arising from selective reporting: positive estimates are preferred in literature. The magnitude of publication bias is sizeable. Hence, the average beauty effect is probably much lower than commonly believed based on the available empirical literature. Taking into account the presence of publication bias, our results do not support the findings provided by Hosoda *et al.* (2003) in the meta-analysis of the beauty effect on job-related outcomes in experimental studies, which imply that beauty is always an asset for individuals.

To determine the key factors that influence the magnitude of the beauty effect and produce heterogeneity of reported results apart from publication bias, we used the Bayesian model averaging technique and OLS-based frequentist check. To check the robustness of our findings, we applied the Frequentist Model Averaging and found that the results are predominantly in line with BMA.

The differences in the reported estimates appear to be driven by sources of real heterogeneity such as individual characteristics of respondents, time spans and geographical regions. Controlling for individual characteristics such as age, education and cognitive skills strongly impact the resulting estimates.

Our results also suggest that study design has an impact on the reported beauty effect in relation to productivity. Researchers who prefer to use beauty ratings based on the self-evaluation of respondents obtain substantially smaller estimates than researchers who use beauty ratings based on the raters evaluation. The authors who choose performance-based measures of productivity, and therefore more homogenous data from specific occupation, obtain higher estimates than those who use earnings as a proxy for measuring productivity. This finding partially explains the large magnitude of beauty effect estimates over the last decade: the most recent studies predominantly examine the effect of physical attractiveness on productivity within occupations.

Another important finding implies that the estimates in our sample do not seem to be significantly different when the occupation requires good looks. This contradicts the reported evidence on higher beauty effect for "dressy" occupations. However, it confirms the findings of Hamermesh & Biddle (1993), who argues that the impact of beauty remains proportional across different types of occupations. We believe that further research is required to provide more evidence on the occupational categorization of beauty in the labor market.

Meta-analysis proves to be an efficient tool when combining the knowledge of empirical literature in the field of beauty economics. We believe that the results of our research will help structure the evidence of the beauty effect in relation to labor productivity. Furthermore, it will increase the understanding of the role of beauty in the labor market.

Bibliography

Ahn, S. C. & Y. H. Lee (2013): "Beauty and Productivity: the case of he ladies professional golf association." *Contemporary Economic Policy* **32(1)**: pp. 155–168.

Andreoni, J. & R. Petrie (2008): "Beauty, gender and stereotypes: Evidence from laboratory experiments." *Journal of Economic Psychology* **29(1)**: pp. 73–93.

Arunachalam, R. & M. Shah (2012): "The prostitute's allure: The return to beauty in commercial sex work." *The B.E. Journal of Economic Analysis & Policy* **12(1)**.

Bakkenbuell, L.-B. (2017): "The impact of attractiveness on athletic performance of tennis players." *International Journal of Social Science Studies* **5(3)**: p. 12.

Berggren, N., H. Jordahl, & P. Poutvaara (2010): "The looks of a winner: Beauty and electoral success." *Journal of Public Economics* **94(1-2)**: pp. 8–15.

Berri, D. J., R. Simmons, J. V. Gilder, & L. O'Neill (2010): "What does it mean to find the Face of the Franchise? Physical Attractiveness and the Evaluation of Athletic Performance." *Working Papers 611506*, Lancaster University Management School, Economics Department.

Biddle, J. & D. Hamermesh (1995): "Beauty, Productivity and Discrimination: Lawyers Looks and Lucre." *NBER Working Papers 5366*, National Bureau of Economic Research, Inc.

Boo, F., M. Rossi, & S. Urzua (2013): "The labor market return to an attractive face: Evidence from a field experiment." *Economics Letters* **118(1)**: pp. 170–172.

Cipriani, G. P. & A. Zago (2011): "Productivity or Discrimination? Beauty and the Exams." *Oxford Bulletin of Economics and Statistics* **73(3)**: pp. 428–447.

Deryugina, T. & O. Shurchkov (2015): "Now you see it, now you do not: The vanishing beauty premium." *Journal of Economic Behavior and Organization* **116**: pp. 331–345.

Doorley, K. & E. Sierminska (2012): "Myth or Fact? The Beauty Premium across the Wage Distribution." *IZA Discussion Papers 6674*, Institute for the Study of Labor (IZA).

Dossinger, K., C. R. Wanberg, Y. Choi, & L. M. Leslie (2019): "The beauty premium: The role of organizational sponsorship in the relationship between physical attractiveness and early career salaries." *Journal of Vocational Behavior* **112**: pp. 109–121.

Doucouliagos, C. (2011): "How large is large? preliminary and relative guidelines for interpreting partial correlations in economics." *Working papers*, Deakin University, Department of Economics.

Doucouliagos, H. & T. D. Stanley (2009): "Publication selection bias in minimum-wage research? a meta-regression analysis." *British Journal of Industrial Relations* **47(2)**: pp. 406–428.

Efendic, A., G. Pugh, & N. Adnett (2011): "Institutions and economic performance: A meta-regression analysis." *European Journal of Political Economy* **27(3)**: pp. 586–599.

Eicher, T. S., C. Papageorgiou, & A. E. Raftery (2009): "Default priors and predictive performance in bayesian model averaging, with application to growth determinants." *Journal of Applied Econometrics* **26(1)**: pp. 30–55.

Fisher, R. A. (1954): "Statistical methods for research workers." *Hafner Publishing Company,New York.* **13(6)**: pp. 429–430.

Fletcher, J. M. (2009): "Beauty vs. brains: Early labor market outcomes of high school graduates." *Economics Letters* **105(3)**: pp. 321–325.

French, M. T. (2002): "Physical appearance and earnings: further evidence." *Applied Economics* **34(5)**: pp. 569–572.

Frieze, I. H., J. E. Olson, & J. Russell (1991): "Attractiveness and income for men and women in management." *Journal of Applied Social Psychology* **21(13)**: pp. 1039–1057.

Gunby, P., Y. Jin, & W. R. Reed (2017): "Did FDI really cause chinese economic growth? a meta-analysis." *World Development* **90**: pp. 242–255.

Hamermesh, D. & J. Biddle (1993): "Beauty and the Labor Market." *NBER Working Papers 4518*, National Bureau of Economic Research, Inc.

Hamermesh, D., X. Meng, & J. Zhang (2002): "Dress for success–does primping pay?" *Labour Economics* **9(3)**: pp. 361–373.

Hamermesh, D. & A. Parker (2003): "Beauty in the classroom: Professors' pulchritude and putative pedagogical productivity." *Technical report.*

Hamermesh, D. S. (2011): *Beauty Pays: Why Attractive People Are More Successful.* Princeton University Press.

Harper, B. (2000): "Beauty, Stature and the Labour Market: A British Cohort Study." *Oxford Bulletin of Economics and Statistics* **62(s1)**: pp. 771–800.

Havranek, T., R. Horvath, Z. Irsova, & M. RusnĂˇk (2015): "Cross-country heterogeneity in intertemporal substitution." *Journal of International Economics* **96(1)**: pp. 100–118.

Havranek, T., Z. Irsova, & O. Zeynalova (2018): "Tuition fees and university enrolment: A meta-regression analysis." *Oxford Bulletin of Economics and Statistics* **80(6)**: pp. 1145–1184.

Havranek, T., M. Rusnak, & A. Sokolova (2017): "Habit formation in consumption: A meta-analysis." *European Economic Review* **95**: pp. 142–167.

Heilman, M. (1983): *Sex bias in work settings: The lack of fit model*, volume 5. JAI Press.

Heilman, M. & L. Saruwatari (1979): "When beauty is beastly: The effects of appearance and sex on evaluations of job applicants for managerial and nonmanagerial jobs." *Organizational Behavior and Human Decision Processes* **23(3)**: pp. 360–372.

Hernandez-Julian, R. & C. Peters (2017): "Student appearance and academic performance." *Journal of Human Capital* **11(2)**: pp. 247–262.

Hershauer, J. C. & W. A. Ruch (1978): "A worker productivity model and its use at lincoln electric." *Interfaces* **8(3)**: pp. 80–90.

Hosoda, M., E. F.Stone-Romero, & G. Coats (2003): "The Effects Of Physical Attractiveness On Job-Related Ooutcomes: A Meta-Analysis Of Experimental Studies." *Personnel Psychology* **56(2)**: pp. 431–462.

JEFFREYS, H. (1961): *The Theory of Probability (3rd edn).* Oxford Classic Texts in the Physical Sciences, Oxford University Press.

JOHNSON, S. K., K. E. PODRATZ, R. L. DIPBOYE, & E. GIBBONS (2010): "Physical attractiveness biases in ratings of employment suitability: Tracking down the "beauty is beastly" effect." *The Journal of Social Psychology* **150(3)**: pp. 301–318.

KANAZAWA, S. & J. KOVAR (2004): "Why beautiful people are more intelligent." *Intelligence* **32(3)**: pp. 227–243.

KING, A. & A. LEIGH (2007): "Beautiful politicians." *SSRN Electronic Journal* .

KRAFT, P. (2012): *The role of beauty in the labor market.* Ph.D. thesis, Prague: CERGE.

LANGLOIS, J. H., L. KALAKANIS, A. J. RUBENSTEIN, A. LARSON, M. HALLAM, & M. SMOOT (2000): "Maxims or myths of beauty? A meta-analytic and theoretical review." *Psychological Bulletin* **126(3)**: pp. 390–423.

LONG, J. B. D. & K. LANG (1992): "Are all economic hypotheses false?" *Journal of Political Economy* **100(6)**: pp. 1257–1272.

LOUREIRO, P. R. A., A. SACHSIDA, & M. J. C. MENDONCA (2012): "Links between physical appearance and wage discrimination: Further evidence." *International Review of Social Sciences and Humanities, Vol. 1, No. 2, pp. 1-16, 2010. Available at SSRN: https://ssrn.com/abstract=2006648* .

MOBIUS, M. M. & T. S. ROSENBLAT (2006): "Why beauty matters." *American Economic Review* **96(1)**: pp. 222–235.

MOCAN, N. & E. TEKIN (2010): "Ugly criminals." *Review of Economics and Statistics* **92(1)**: pp. 15–30.

Nelson, J. P. (2013): "Meta-analysis of alcohol price and income elasticities – with corrections for publication bias." *Health Economics Review* **3(1)**: p. 17.

Paphawasit and Fidrmuc (2017): "Beautiful Minds: The Impact of Beauty on Research Productivity in Economics." *Working Paper* .

Parrett, M. (2015): "Beauty and the feast: Examining the effect of beauty on earnings using restaurant tipping data." *Journal of Economic Psychology* **49(C)**: pp. 34–46.

Persico, N., A. Postlewaite, & D. Silverman (2003): "The Effect of Adolescent Experience on Labor Market Outcomes: The Case of Height." *PIER Working Paper Archive 03-036*, Penn Institute for Economic Research, Department of Economics, University of Pennsylvania.

Pfann, G., J. Biddle, D. Hamermesh, & C. M. Bosman (2000): "Business success and businesses' beauty capital." *Economics Letters* **67(2)**: pp. 201–207.

Pfeifer, C. (2012): "Physical attractiveness, employment and earnings." *Applied Economics Letters* **19(6)**: pp. 505–510.

Ravina, E. (2008): "Love and loans the effect of beauty and personal characteristics in credit markets."

Ritts, V., M. L. Patterson, & M. E. Tubbs (1992): "Expectations, impressions, and judgments of physically attractive students: A review." *Review of Educational Research* **62(4)**: pp. 413–426.

Sachsida, A., A. C. Dornelles, & C. W. Mesquita (2003): "Beauty and the labor market - study one specific occupation." *SSRN Electronic Journal* Available at SSRN: https://ssrn.com/abstract=505742 or http://dx.doi.org/10.2139/ssrn.505742.

SALTER, S. P., F. G. MIXON, & E. W. KING (2012): "Broker beauty and boon: a study of physical attractiveness and its effect on real estate brokers' income and productivity." *Applied Financial Economics* **22(10)**: pp. 811–825.

SAUERMANN, J. (2016): "Performance measures and worker productivity." *IZA World of Labor* .

SCHOLZ, J. K. & K. SICINSKI (2015): "Facial attractiveness and lifetime earnings: Evidence from a cohort study." *Review of Economics and Statistics* **97(1)**: pp. 14–28.

SEN, A., M.-C. VOIA, & F. R. WOOLLEY (2010): "Hot or Not: How Appearance Affects Earnings and Productivity in Academia." *Carleton Economic Papers 10-07*, Carleton University, Department of Economics.

STANLEY, T. (2005): "Beyond Publication Bias." *Journal of Economic Surveys* **19(3)**: pp. 309–345.

STANLEY, T., H. DOUCOULIAGOS, M. GILES, J. H. HECKEMEYER, R. J. JOHNSTON, P. LAROCHE, J. P. NELSON, M. PALDAM, J. POOT, G. PUGH, R. S. ROSENBERGER, & K. ROST (2013): "Meta-Analysis Of Economics Research Reporting Guidelines." *Journal of Economic Surveys* **27(2)**: pp. 390–394.

STANLEY, T. D. (2001): "Wheat from Chaff: Meta-analysis as Quantitative Literature Review." *Journal of Economic Perspectives* **15(3)**: pp. 131–150.

STANLEY, T. D. (2007): "Meta-regression methods for detecting and estimating empirical effects in the presence of publication selection." *Oxford Bulletin of Economics and Statistics* **0(0)**.

STANLEY, T. D., S. B. JARRELL, & H. DOUCOULIAGOS (2010): "Could it be better to discard 90% of the data? a statistical paradox." *The American Statistician* **64(1)**: pp. 70–77.

Talamas, S. N., K. I. Mavor, & D. I. Perrett (2016): "Blinded by beauty: Attractiveness bias and accurate perceptions of academic performance." *PLOS ONE* **11(2)**.

Valickova, P., T. Havranek, & R. Horvath (2014): "FINANCIAL DEVELOPMENT AND ECONOMIC GROWTH: A META-ANALYSIS." *Journal of Economic Surveys* **29(3)**: pp. 506–526.

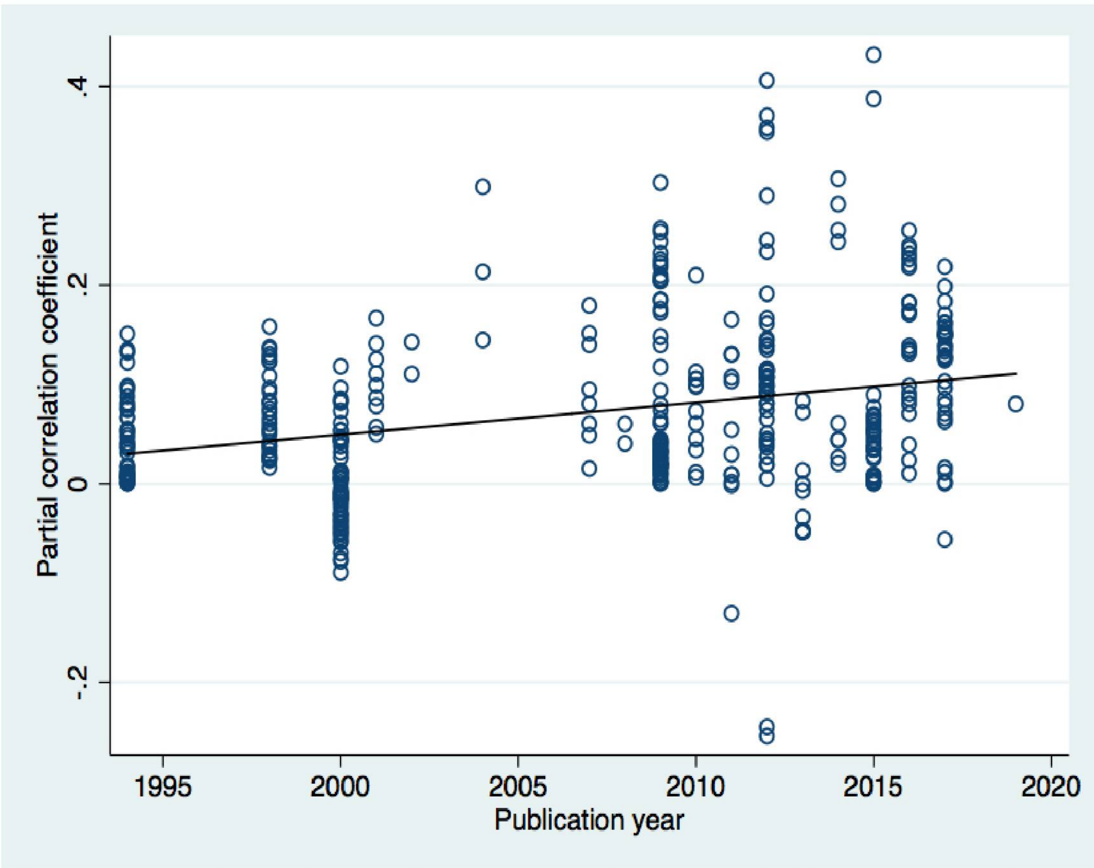
Von Bose, C. (2012): "Child stars vs. ugly ducklings: Does adolescent attractiveness contribute to the beauty premium." *Technical report*, Technical report, Working paper.

Wilson, R. K. & C. C. Eckel (2006): "Judging a book by its cover: Beauty and expectations in the trust game." *Political Research Quarterly* **59(2)**: pp. 189–202.

Appendix A

Beauty effect estimates over time

Figure A.1: Beauty effect estimates over time

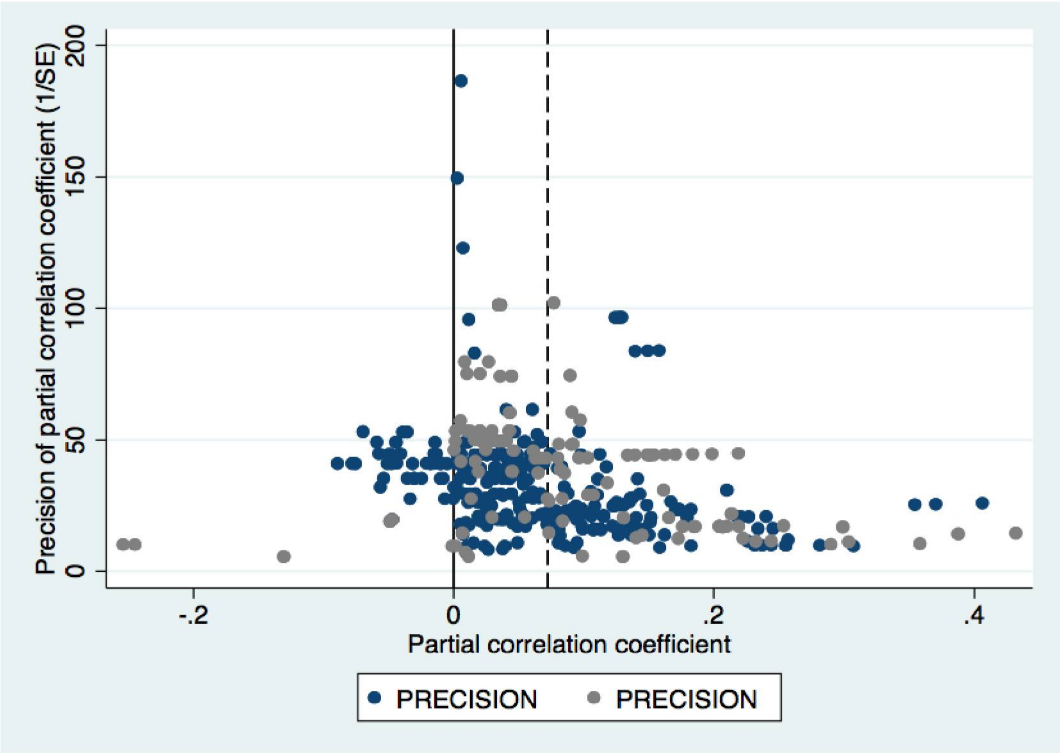


Notes: The figure shows distribution of partial correlation coefficients of the reported beauty effect over time. The time trend is not statistically significant

Appendix B

Funnel plot of beauty effect estimates by productivity category

Figure B.1: Funnel plot of beauty effect estimates by productivity category

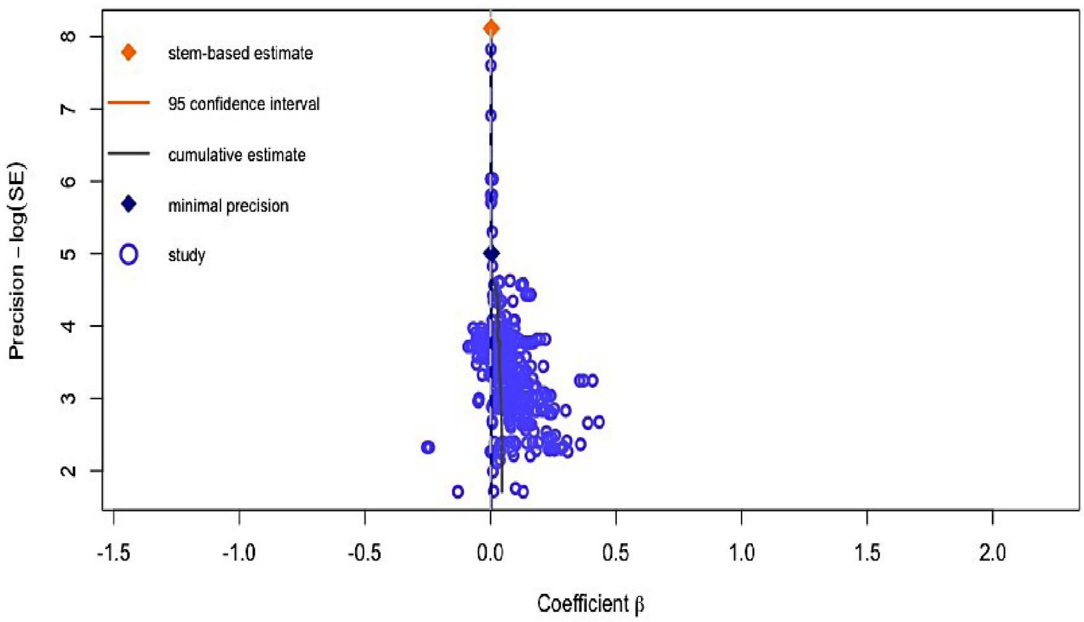


Notes: The earning-based beauty effects are represented in blue color, the performance-based-in grey color. The dashed vertical line demonstrates the mean partial correlation coefficient.

Appendix C

Stem Approach Visualization

Figure C.1: Stem approach visualization



Notes: The figure illustrates a funnel plot of stem-based bias correction method. The y-axis denotes a measure of precision (logarithm of standard error of the beauty effect estimate). The orange diamond at the top indicates the stem-based estimate. The navy diamond at the middle of the plot demonstrates the minimal level of precision for the inclusion. The stem-based estimate is defined by the studies, whose precision are above the navy diamond.

Appendix D

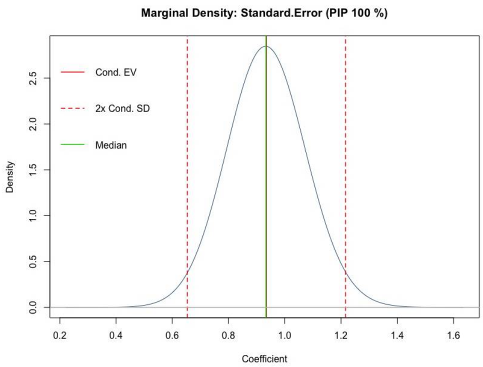
Diagnostics of BMA

Table D.1: BMA Summary

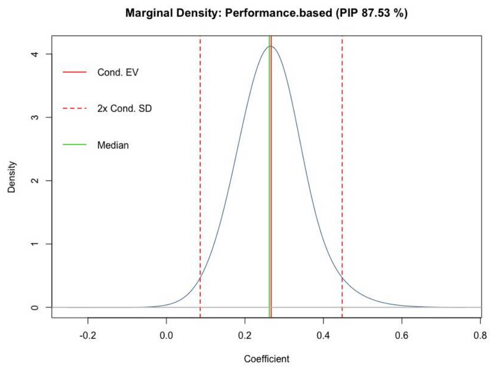
Mean no. regressors	Draws	Burnins
6.5176	2e+06	1e+06
Time	**No. models visited**	**Modelspace 2K**
4.090759 mins	632126	524288
Percent visited	**Percent Topmodels**	**Corr PMP**
121	99	0.9998
No. Obs.	**Model Prior and uniform**	**Percent g-Prior**
418	9.5	UIP

Notes: The table reports summary of Bayesian Model Averaging estimation. We use the the uniform model prior and the unit information prior. The results of this BMA procedure are reported in Table 4.5. UIP=Unit Information Prior; PMP= Posterior Model Probablity

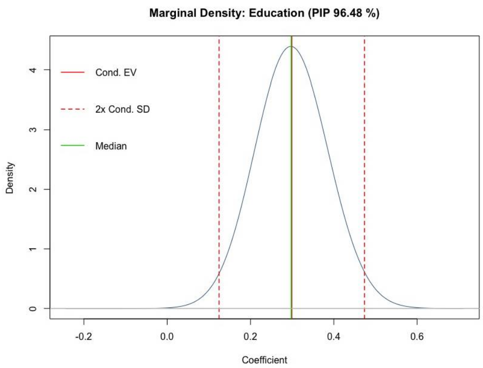
Figure D.1: Posterior coefficient distributions for the most important factors



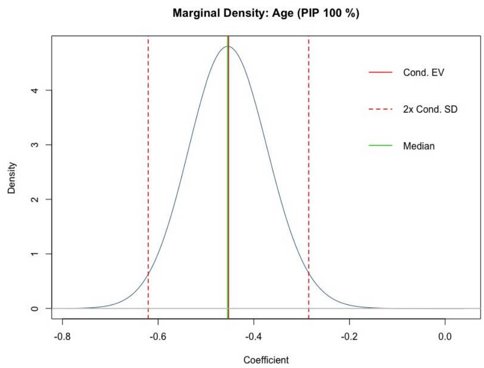
(a) Standard Error



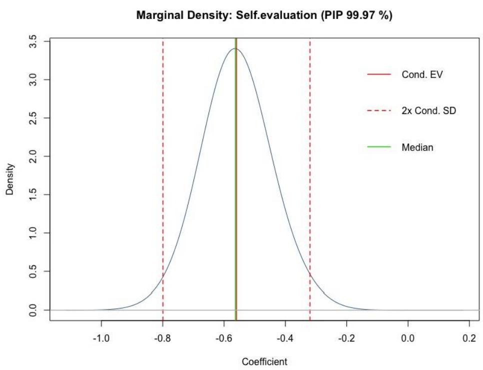
(b) Performance Based Productivity



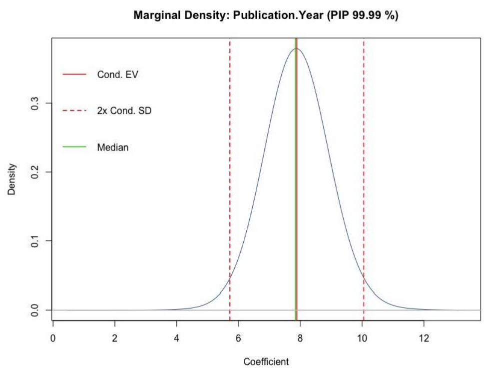
(c) Education



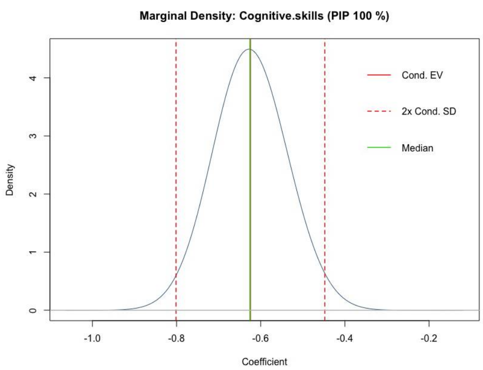
(d) Age



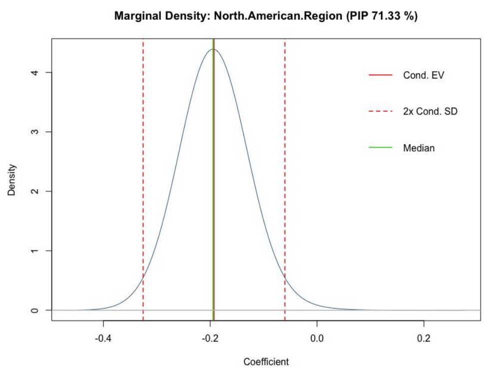
(e) Self-Evaluated Beauty



(f) Publication Year



(g) Cognitive Skills



(h) North American Region

Notes: The figure represents the densities of the estimation parameters from Table 4.5 with the highest PIP.