CEP 10-07

Hot or Not: How appearance affects earnings and productivity in academia

Anindya Sen
University of Waterloo

Marcel Voia
Carleton University

Francis Woolley
Carleton University

October 2010

CARLETON ECONOMIC PAPERS

Department of Economics
1125 Colonel By Drive
Ottawa, Ontario, Canada
K1S 5B6

The effect of hotness on pay and productivity

October, 2010

Anindya Sen, Marcel Voia and Frances Woolley

Contact information:

Anindya Sen: Department of Economics, University of Waterloo, Hagey Hall of Humanities, Waterloo, Ontario, Canada N2L 3G1, email: asen@uwaterloo.ca voice: 519 888 4567 x33695 fax: 519 725 0530

Marcel Voia: Department of Economics, Carleton University, 1125 Colonel By Drive, Ottawa, Ontario K1S 5B6 email: marcel-cristian_voia@carleton.ca voice: 613-520-2600 x 3546 fax: 613 520 3906

Frances Woolley: Department of Economics, Carleton University, 1125 Colonel By Drive, Ottawa, Ontario K1S 5B6 email: frances_woolley@carleton.ca voice: 613 520 2600 x 3756 fax: 613 520 3906

Acknowledgements: We would like to thank Hideki Ariizumi and Natasha De Sousa for their contribution to data gathering, Martin Dooley and Elaine Constant for helping us track down rank data, and Kim Lehrer for excellent comments at the CEA meetings in Quebec City.

Abstract: In this paper we examine the impact of a professor's hotness, as rated by students, on his or her salary, controlling for research and teaching productivity. We also estimate the impacts of a professor's hotness on the quality of his or her teaching, as evaluated by students, and the impact of hotness on research productivity, as measured by citations, publications, co-authorship, and grant funding. Our study is based on data describing economics professors at sixteen universities. Although a relatively small proportion of our sample is rated "hot" by students, hotness generates, for some, a significant earnings premium, even with comprehensive controls for productivity. We find a strong relationship between hotness and teaching productivity, but a much weaker relationship between hotness and research productivity. The unique contribution of this paper is the use of data on actual productivity, which is generally unavailable in papers assessing the returns to appearance.

The effect of hotness on pay and productivity

Our paper is motivated, first, by a long-standing puzzle: Are beautiful people paid more because they are more productive, or because employers or consumers have a taste for beauty? (Hamermesh and Biddle, 1994; Mobius and Rosenblat, 2006; Kuhn and Shen, 2009) Our second motivation is the availability of data. The website ratemyprofessors.com allows students to rate their professors' appearance: "(just for fun): __ hot __ not." In this paper we employ a unique panel data set with information on academic salaries, a wide range of measures of both the quantity and quality of research and teaching output, extensive human capital controls, and our key explanatory variable of interest, hotness. The contribution of our paper is to estimate the impact of hotness on earnings controlling for productivity, *and* the impact of hotness on direct measures of productivity, namely teaching and research output. We are not aware of any other paper that is able to use real-world data on productivity when estimating the economic impacts of attractiveness.

Although ratemyprofessors asks students to evaluate "appearance," the framing used, hot or not, has sexual connotations. Hotness captures both beauty *and* the multiple dimensions of personality that contribute to sexual attractiveness. Understanding this is crucial to the interpretation of our findings. The *absence* of a return to hotness strongly suggests that beauty has no real effects. The *presence* of a positive return for hotness, however, may reflect either returns to beauty or returns to sexual attractiveness. We interpret our findings, therefore, in the context of the growing body of research on the relationship between attractiveness, masculinity/femininity, and labor market outcomes (Johnson and Tassinary, 2007; Bowles, Babcock and Lai, 2007). Our finding that, for certain groups, hotness has a significant effect on pay contributes to the growing body of evidence that personality (Nyhus and Pons, 2005) and sexuality (Miller, Tybur, and Jordan, 2007) matters in the labor market.

Hotness information was included in the ratemyprofessors.com website for fun, but our analysis uses serious methods. Panel data and an extensive list of covariates allow rich controls for individual heterogeneity. We perform numerous robustness tests, and find that our results, particularly for the effect of hotness on salaries, hold across a wide range of estimation techniques and specifications. Our results documenting which demographic groups receive a premium for hotness, and which measures of productivity respond most to hotness, enhance understanding of the academic labor market.

1. Related Literature

In this paper we draw upon two related literatures. The first is research on academic labor markets, for example, studies of gender and other determinants of pay (Barbezat, 1987; 1991), promotion (Ginther and Kahn, 2004), co-authorship (Boschini and Sjögren, 2007) and citation rates (Di Vaio, Waldenström and Weisdorf, 2009). Findings that are particularly useful for interpreting our results are Blackaby, Booth and Frank's (2005) result that the ability to generate outside offers is an important source of academic salary variation, and Warman, Worswick and Woolley's (2010) demonstration of the cohort effects in academic salary determination.

We also draw upon the small existing literature on professorial pulchritude. A number of studies have found a positive relationship between instructors' attractiveness and teaching evaluations using either independent measures of instructors' beauty (Hamermesh and Parker, 2005; Süssmuth, 2006) or ratemyprofessors.com hotness data (Felton, Mitchell and Stinson, 2004; Riniolo, Johnson, Sherman, and Misso, 2006; Freng and Webber, 2009). With respect to other dimensions of academic success, Hamermesh (2006) finds that better-looking economists are more likely to be elected to the American

Economics Association executive, while Green, Mixon , and Trevino (2005) find empirical support for the hypothesis that, because attractiveness increases teaching productivity more than it increases research productivity, hot professors would be more likely to chose careers at liberal arts universities. However we know of no other studies that link hotness to earnings and research productivity.

2. Data

The population analyzed in this paper is all tenured and tenure-track faculty members appointed to the sixteen university economics departments in the province of Ontario, Canada, at some point between 1996 and 2006, with the exception of those holding university-level administrative positions. This population was identified through individual and university web pages, university calendars, and the Canadian Economics Association newsletter. Data were obtained on individuals' salaries and productivity from 1996 to 2006 using a variety of publicly available sources such as provincial government web sites, granting agencies' competition results, the Econlit database and the Social Sciences Citation Index. The resulting panel comprises 493 individuals, observed over time to create a sample size of 3,455 observations. Table A1 provides descriptive statistics, while Appendix B provides a detailed summary of the sources used for each variable. Of these various data sources, two - those for salary information and for hotness -- merit detailed discussion.

Our salary data is obtained from public sector salary disclosure information published by the Ontario Ministry of Finance. As a result of the 1996 Ontario Public Sector Salary Disclosure Act, the salaries of all university professors earning $100,000 or more are listed. The salary release consists of an individual's calendar year (January to December) gross earnings.[1] In 1996, 13.2 percent of the

[1] Additional duties carried out through an individual's own university (for example, teaching a summer course) are included; outside consulting or other activities (for example, teaching at another university while on sabbatical) are excluded. Individuals on maternity or parental leave receive additional government payments which are not

individuals in our sample had salaries over $100,000, but by 2006 we are able to observe directly the salaries of 58.2 percent of economics professors.

Hotness data is obtained from the website www.ratemyprofessors.com (RMP). On this site, students anonymously evaluate their professors' easiness, helpfulness, clarity and "appearance". Helpfulness and clarity are averaged to generate a single rating for quality, while ease and appearance are reported separately. A sample rate my professor form is provided in Appendix A. We utilize responses to the question: "Appearance (just for fun): __ hot __ not." Ratemyprofessors assigns a score of +1 for each hot rating and -1 for each not. If the number of hots exceeds the number of nots, a chili pepper appears in the professors' summary.[2] Since our attempts to obtain historical data through web archives or directly from ratemyprofessors.com were unsuccessful, we use the cumulative ratings between April 21, 2001, when ratemyprofessors added Canadian schools[3], and the time that we collected the data in 2009.

There are some limitations to ratemyprofessors data. Students non-randomly self-select onto ratemyprofessors.com. Intuition suggests that students who particularly like or dislike a professor are more likely to enter their ratings on the website relative to individuals who are less pronounced in their tastes. The number of ratings provided is not large. In our sample, the mean number of ratings is 16 (see Table A1) including those who receive no ratings – a group that consists primarily of older faculty and those teaching small upper level or graduate courses. Non-students can create ratings, and a single student can create several positive or negative ratings. Yet official teaching evaluations also suffer from

---

included here, however we control (imperfectly) for this with the generated variable 'sabbatical,' a dummy that indicates the individual has experienced a year-on-year drop in salary.

[2] http://www.ratemyprofessors.com/faq.jsp

[3] http://web.archive.org/web/20010622210910/www.ratemyprofessors.ca/canada/New.asp

selection bias, for example, when in-class evaluations are used, students who drop the course, do not attend class, or simply cannot be bothered to fill out the form are not included.

There are several arguments that can be made in support of using ratemyprofessors ratings. First, they correlate highly with official university evaluations (Timmerman, 2008: abstract), which suggests that the selection bias on ratemyprofessors is no worse than selection bias in other forms of student-based evaluation. Hoopes and Albrecht (2007) find that "the rankings of official university evaluations and RMP [ratemyprofessors.com] are very similar…lending validity to RMP as a tool for intra-university comparison of potential professors." Second, ratemyprofessors is written by students for students, and is often the only source of information available about the quality of a potential instructor. In the course of writing this paper, we read hundreds of evaluations. Overwhelmingly, the ratings simply try to provide useful information for other students, for example, "You need to listen to him carefully by sitting in the very front. Otherwise you will do very bad in his class."[4]

A third argument in support of using ratemyprofessors rankings is that openness creates quality controls. Misleading ratings can be and are corrected by subsequent users, as in "Best econ prof at ______, ignore the bad reviews."[5] This is particularly true of our hotness variable, where a spurious hot vote is negated by a single not vote. Fourth, users may flag offensive ratings, and ask that they be removed. Finally, the ratemyprofessor rankings reflect, on average, 16 evaluations of a professor's hotness. This contrasts with, for example, Hamermesh and Parker (2005), for example, which base beauty estimates on six evaluations. Moreover, students rate a professor on how attractive he or she is in real life – a rating that includes the person's voice, gestures, and so on. Hence it provides a better

[4] Rating of Professor ___________, University of Ottawa, source: www.ratemyprofessors.ca.

[5] Rating of Frances Woolley, source: www.ratemyprofessors.ca, accessed July 23, 2010

estimate of the effect of real world attractiveness than commonly used measures such as ratings of photographs.

There are some limitations to our hotness measure. For example, to test the hypothesis that the beautiful are paid more because of employer discrimination, we would ideally like to know how beautiful each individual in our sample was when, say, they were hired. Unfortunately, we observe individuals at different stages in their life-cycle, and there is a strong correlation between perceived hotness and age.[6] For example, in our sample, 27 percent of female assistant professors are hot, but no female full professor has a chili pepper. For men, 20 percent of assistant professors are hot, compared to six percent of full professors. We overcome this limitation through the use of age controls, by breaking down our sample by gender and cohort, drawing upon individual student comments and integrating our results with the broader literature on physical attractiveness.

3. The Effect of Hotness on Pay: Empirical strategy

Our model is straightforward adaption of Hamermesh and Biddle (1994). Faculty members' earnings depend upon appearance and productivity:

$$Y_{it}^{*} = \beta_a \vartheta_i + \alpha_0 + \alpha_i + \alpha_t + \alpha_m + X'_{it,P}\beta_P + X'_{it,NP}\,\beta_{NP} + X'_{it,T}\,\beta_T + \varepsilon_{it} \qquad (1)$$

Equation (1) is a simple human capital earnings model, augmented to include hotness. The variable Y_{it}^{*} is individual i's earnings at time *t*. We use the notation Y_{it}^{*} to indicate that salary is a latent variable, that is, we do not observe every individual's actual salary, as discussed below. Warman, Worswick, and Woolley (2010) find that academic salaries in Canada follow a normal distribution. Accordingly, we use salaries, as opposed to log salaries, as our dependent variable. We adjust all salary data for inflation using the Consumer Price Index (CPI). The error term, ε_{it} , is assumed to follow a normal distribution,

[6] The correlation between "hot" and year of PhD is 0.22 – note that year of PhD is decreasing in age.

with mean zero. Since we observe the same individuals repeatedly, the error terms are not independent, but instead are clustered at the individual level.

The variable ϑ_i is a time-invariant measure of an individual's hotness, taking on a value of one if the number of hot ratings on ratemyprofessors.com exceeds the number of not ratings – that is, if the professor has a chili pepper. Our study differs from others such as Hamermesh and Biddle (1994) in that we consider just one occupation. In a single occupation hotness could be *negatively* correlated with earnings. If there is a bias against unattractive individuals at the time of hiring then, all else being equal, a less attractive individual must be better qualified in other respects to obtain an academic position (have a more interesting job market paper, for example). If these other characteristics are more closely correlated with long-run academic success, those hired for looks will eventually fall behind. Morever, if Green, Mixon, and Trevino's (2005) hypothesis that looks are valued more at teaching institutions is correct, the concentration of the more attractive at less research-intensive universities could generate an inverse relationship between hotness and salary. Finally, academic salaries increase with age (Warman, Woolley, and Worswick, 2010), but hotness does not. This could lead to an inverse appearance/earnings relationship if we have imperfect controls for age.

Apart from hotness, equation (1) is a standard earnings regression. The term α_0 is a constant. The α_i are dummy variables for the 16 universities in our sample, and capture university-level salary differentials. The α_t are time dummies that capture year-to-year variation in real salaries. We include the dummy variable α_m to capture gender pay differentials. Together these sets of dummies control for the effect of non-productivity related determinants of salaries such as local living costs, overall university funding levels (Martinello 2006) and gender.

The covariates $X'_{it,P}$ in equation (1) are productivity controls. Earnings regressions are typically based on proxies for productivity, for example, experience and education. We use years since PhD and its squared term to measure experience. Educational quality is measured indirectly through controls for having an American PhD or other foreign PhD. What is highly unusual about our study is that we measure research productivity directly. Life-time productivity is captured using controls for rank (assistant, full, other, with associate being the omitted category). Productivity in the recent past is captured with controls for the number of journal publications in the Econlit database (and publications squared),[7] the number of books listed in Econlit, the number of citations in the social sciences citation index database, and amount of Social Sciences and Humanities Research Council (SSHRC)[8] research grant funding obtained in the previous year. We use one year lagged values of these productivity measures because, in all the universities in our sample, a faculty member's annual salary increase is based on collectively negotiated increments and, if the university has merit pay, on productivity in the past year (Chant, 2005). Citations, however, are defined as citations of all works published over a life-time, and SSHRC grant funding is based on productivity in the five years prior to the grant application date, so although these are annual variables, they reflect productivity over a longer time span.

The $X'_{it,NP}$ in equation (1) are covariates that are non-productivity specific. We construct dummy variables that capture whether or not an individual is currently on sabbatical, whether or not the individual is a department chair, and the type of remuneration scheme employed by the university. Using Chant's (2005) typology, we divide the universities in our sample into those with and without

[7] We also experimented with including controls for publication quality. Adding a control for the presence of a top 10 publication in the previous year increased the estimated marginal effect of hotness in Table 1 case 5 to from 12,957, significant at p=0.01, to 13,465, also significant at p=0.01.

[8] The Social Sciences and Humanities Research Council of Canada is the primary source of research funding for most economists in the country. SSHRC grants are highly competitive, and the awards are determined by a committee comprised of academic economists.

unions, merit pay[9] and salary caps. We include dummy variables to control for the type of pay scheme present at the university.

The variables $X'_{it,T}$ are controls for the quantity (number of ratings) and quality (helpfulness, clarity and ease) of teaching. Our measure of teaching productivity is, however, a mix of past, present and future output, consisting of average ratemyprofessor scores based on cumulative ratings since 2001. We include ε_{it} as the time-varying error component.

In estimating equation (1), we face both selection and censoring issues. The selection issue arises because not all individuals are rated on ratemyprofessors, and the selection of individuals to be rated is non-random. Older professors, for example, are less likely to be rated. We address this selection issue by using an observed variable, a dummy equal to one for professors who are not rated, that identifies the unrated professors and proxies this selection into rating. This selection variable is consistently important in explaining academic pay. We also model selection into the ratemyprofessors database, as described in Appendix C, and find that, if anything, the impact of hotness on pay and productivity is amplified when we take the potential endogeneity of being rated into account.

The censoring issue arises because, since our data is obtained from the Ontario salary disclosure, we have salary information for those earning over $100,000 only. This is a standard "type 1" tobit situation (Amemiya, 1985). Our observed variable, Y_{it}, is equal to the true salary, Y_{it}^*, for individuals earning over $100,000; otherwise, we observe only the lower limit of $100,000.

$$Y_{it} = Y_{it}^* \text{ if } Y_{it}^* > \$100{,}000,\ \$100{,}000 \text{ otherwise} \qquad (2)$$

[9] In a number of Ontario universities, there is no merit pay – an individual's annual increment is independent of his or her teaching or research productivity (Chant, 2005). In universities with salary caps, any individual earning at or above the cap can only receive a salary increase if the cap increases.

Combining equations (1) and (2), our dependent variable is defined as:

$$Y_{it} = \begin{cases} \beta_a \vartheta_i + \alpha_0 + \alpha_i + \alpha_t + \alpha_m + X'_{it,P}\beta_P + X'_{it,NP}\,\beta_{NP} + X'_{it,T}\,\beta_T + \varepsilon_{it} \text{ if } Y_{it}^{*} > \$100,000 \\ 100,000 \text{ otherwise} \end{cases} \quad (3)$$

We estimate equation (3) using a standard Stata maximum likelihood tobit regression that accounts for individual and group specific effects as well as time effects. As the data used to estimate this model are observed over a period of 11 years (from 1996 to 2006), and individuals are hired and retire during this period, we have an unbalanced panel. We recognize the panel nature of our data by clustering the standard errors by individual in order to account for unobserved correlations across individuals over time.

Because of inflation, equation (3) is not an exact description of our data. For example, in the first year of our data, the salary cut-off was $100,000 in 1996 dollars, which is more than $100,000 in 2006 dollars. Unfortunately it is not possible to control for year-to-year censoring and still maintain the panel aspects of our data, but the robustness checks reported in Appendix Table A2 shows that our results still hold with different inflation adjustments. We also report the results of different empirical specifications (e.g. probit, linear regression on those earning over $100,000 only) and find that our results are robust to our modeling choices.[10]

The key variable of interest – hotness – is time invariant, and we identify its impact through variations in hotness across individuals. This raises the possibility that coefficient estimates of hotness might be confounded by unobserved heterogeneity specific to each individual. As outlined in the discussion of equation (1) above, we use a rich set of covariates to control for any individual heterogeneity that is associated with an observable variable. For example, university fixed effects

[10] We obtain similar results using 2006 only (though significance decreases to the 10% level due to the decrease in sample size), nominal dollars (i.e. no adjustment for inflation) or 1996 dollars (with a lower limit of $100,000 in 1996 dollars).

capture individual heterogeneity arising from differences in pay scales across universities and differences in the emphasis placed on teaching versus research.

Another possible concern is simultaneity bias between earnings and hotness. In our analysis, we assume that variation in hotness across individuals is exogenously correlated with earnings. However, it is possible that an increase in earnings allows individuals to invest in appearance, by buying nicer clothes, for example. If this is a strong relationship then single equation coefficient estimates of hotness will be correlated with the error term. We have no instrument to correct for this type of bias. However, while falling short of a fully satisfactory 'fix', the use of university and year dummies should purge some of this bias if it is related to location or year specific shocks.

4. Do hot professors earn more?

Table 1 shows the results of the estimation of equation (3) above, the impact of appearance on earnings. Earnings are measured as annual salary in 2006 dollars, adjusted for inflation using the Consumer Price Index. Hot is a dummy variable taking on a value of one if the professor had a chili professor, or more hots than nots, on ratemyprofessors.com. Table 1 includes asterisks noting where we have controlled for other determinants of salary such as research productivity.

Table 1 shows that a hot appearance has a significant positive effect on salary. Table 1 reports three sets of the marginal effects. The first corresponds to the coefficient β_a in equation (3), and takes into account the full effect of hotness, that is, the effect of being hot *given* that the individual is earning over \$100,000, and also the effect of being hot *on the probability* of earning over \$100,000. The effect of hotness on salary is, for our sample as a whole, always positive. Once controls for differences between universities, changes over time and productivity are added, the effects are both economically and statistically significant. In the specification with the best fit, Case 5, the coefficient on appearance

($12,957) is somewhat larger than earnings premium enjoyed by full professors over associates ($9,626, s.e. $2,598) and similar in size to the male earnings premium($11,423, s.e. $3296). To put this number into perspective, bear in mind that, conditional upon earning $100,000 or more, the average salary in our sample is $127,017, so our highest estimates of the effects of hotness on earnings are approximately 10 percent of the average observed salary.

More insight into the effect of appearance on earnings can be gained by computing two further marginal effects. The marginal effect of hotness conditional on being uncensored gives the earnings premium associated with being hot for those earning $100,000 or more. Among this group, the hot earn between $1,785 to $3,241 more than the not, with the larger (and statistically significant) estimates corresponding to specifications that control for years since PhD and years since PhD squared. The final marginal effect shows the impact of being hot on the probability of earning over $100,000. The hot are between 6 and 17 percentage points more likely to earn over $100,000. Again, the larger and statistically significant estimates correspond to specifications with controls for years since PhD. These results suggest that the large marginal effects reported in the first line of Table 1 are driven primarily by the greater likelihood of earning over $100,000 if an individual is hot – the actual effect of hotness on those already earning over $100,000 is considerably smaller.

The effects reported in Table 1 seem counter to Kuhn and Shen's (2009) argument that preferences for beauty are stronger in occupations with relatively lower skill requirements. Since university professors must have high levels of education, and research and teaching are skill-intensive, we would not expect hotness to have a large effect on academic salaries. It could be that hotness is a proxy for valuable traits such as confidence (Mobius and Rosenblat, 2006), professional ambition that motivates investments in appearance, or for health and reproductive fitness (Jason Weeden and John Sabini, 2005). Yet while hotness may be affecting salary indirectly through the productivity-enhancing effects of confidence, ambition or health, it should be emphasized that hotness still has an effect on earnings even with rich controls for productivity.

Further insight into the hotness/earnings relationship can be gained by breaking up our group into smaller sub-samples. This is done in Table 2. We use the specification corresponding to "Case 5" in Table 1 because it has the highest log likelihood score, so provides the best fit. The first two columns of Table 2 show that the positive effect of appearance on earnings is driven entirely by males. In column 2 we find an insignificant negative relationship between hotness and pay for women.

There are a number of reasons why men and women might receive differential returns to hotness. For example, hot female professors' comments often refer to their physical appearance:

> She's a "talk & chalk" type of professor, and it worked well with the subject matter for ECO___. She keeps it simple. If you go to class, take decent notes, and have common sense, you will do well. Also, she's a fox. Which is a nice added bonus." "Genuinely nice person. She is also dreamy." "Took the class with her a few years ago as an elective, and got 99 with minimal effort. She was a pretty good professor from what I remember - she actually cares about the students. She's also easy on the eyes, so it's worth going to class." "I couldn't stand her in the end despite she is good looking."

These comments evaluate teaching --"pretty good professor" –and appearance --"also easy on the eyes" separately. Men receive personal comments also, for example "mummmmmmmmmm" , "He's so hot" or "Looks a bit like Tim Robbins", but such comments on a man's appearance are rare. A more typical review is "Great Prof. Very funny. Very helpful. Tough assignments prepare you for challenging (but fair) exams. You actually LEARN in this class.. but do NOT miss any lectures." Women might not experience the same salary premium for being hot if their hotness is more strongly determined by physical beauty.

As noted in the empirical strategy section above, within academia the predicted effect of beauty on salary is ambiguous, since an attractive appearance may allow otherwise less productive people to get hired.

Further explanations of why hotness might not pay for women come from economic psychology. Masculine men and feminine women are more likely to deemed attractive than those of more androgynous appearance (Johnson and Tassinary, 2007). Other research suggests that when women are more demanding in negotiations, they are rated as less attractive (Bowles, Babcock and Lai, 2007). It could be the case that characteristics that enable women to negotiate higher salaries also reduce women's attractiveness in the eyes of their students. Other studies (Nyhus and Pons, 2005) have found that having an "agreeable" personality reduces women's earnings. If, for women, hotness is associated with being agreeable ("Genuinely nice person"), this would provide another explanation of why hotness might not pay.

The age/hotness relationship is also revealing. The last four columns of Table 2 shows how the estimated effect of hotness changes as we gradually expand our sample from the youngest academics – those who received their PhD after 2000 – and include a larger and larger portion of the overall sample. The absence of a significant hotness return for those who earned a PhD after 1995 may reflect, in part, the limitations of our data. Since we only have information on individuals earning over $100,000, we do not observe salary data for some of the younger professors. Yet the finding of no return to hotness is consistent with the hypothesis that, conditional upon receiving an academic job, there is no discrimination in favor of the beautiful.

The effect of hotness on earnings increases substantially when we include mid-career academics. The hotness premium is primarily driven by those who received their PhD after 1985, so were in their mid 30s or older when the hotness data began to be collected. We believe the for older professors reflect students' demanding standards for hotness. A high quality not-hot professor will typically be described as a "Great prof" with the occasional "awesome." By way of contrast, these are typical reviews for senior professors who are rated as hot:

> "I absolutely loved Professor ____. He's one of the busiest people in the Economics department and yet he never sees it as an inconvenience to do what he can to help." "I absolutely love Professor _____. He is amazing and the best prof at _____and in Canada!!!"

A professor has to be special to earn a chili pepper if he's over 40 -- "a semi-retired superman" as one hot professor is described – charismatic, charming, well-organized and likeable.

Why might semi-retired supermen earn more? In most Ontario universities, an academic negotiates his or her initial starting salary, and from that point onwards pay increases are largely determined by standard increments built into the collective agreement (Chant, 2005; Warman, Woolley and Worswick, 2010). The best way to negotiate a substantial pay increase is to obtain an outside offer. Indeed, Blackaby, Booth and Frank (2005) find that the ability to generate outside offers is a key factor explaining earnings inequality among academic economists in the UK. Although we cannot test this directly, it seems plausible that charismatic and likeable professors might be better able to generate outside offers. This also would explain why we find a hotness effect for men only – Blackaby, Booth and Frank (2005) found that women typically had difficulty generating outside offers because they were less mobile.

5. Teaching productivity

The findings in Tables 1 and 2 suggest that men, in particular, are paid more if they have a hot appearance. But are the hot also more productive? Most papers in the beauty literature are unable to answer this question, as they lack direct measures of productivity. For academics, however, productivity can be measured directly by counting publications, research grants, citations, and so on. In the remainder of the paper, we explore the effect of hotness on productivity. It should be noted that the effects of hotness on productivity found in this section do not explain the premiums for hotness reported in Tables 1 and 2 as these regressions included controls for productivity. Instead, the effects of hotness on productivity constitute additional indirect impacts of hotness on earnings.

There are two components of academic productivity, teaching and research. We measure teaching productivity using students' perceptions of instructor quality, Q_i, as measured by instructors' scores on the ratemyprofessor website. (Unfortunately we have no other indicators of teaching productivity, such as the amount learnt by students.) Our hypothesis is that perceived teaching quality is a function of hotness, experience and teaching effort and research productivity. Hotness can affect measured teaching quality in a number of ways. First, consumers (students) may discriminate in favour of more attractive professors, giving them higher rankings for the same level of teaching quality. Second, more attractive people may indeed be better at conveying information, perhaps because students pay more attention ("She's also easy on the eyes, so it's worth going to class.") Third, ratemyprofessors hotness rankings may proxy underlying personality traits, for example, charisma or empathy, which are valuable in teaching. To clarify the relationship between hotness and teaching productivity, we also consider the effect of hotness on students' ratings of professors' easiness.

Ratemyprofessors scores have four components: helpfulness, clarity, quality (a weighted average of helpfulness and clarity), and easiness. All are measured on a scale of one (which ratemyprofessors.com describes as useless, incomprehensible, hard) to five (or, in ratemyprofessors.com language, extremely helpful, crystal, easy). We assume

$$Q_i = \beta_a \vartheta_i + \alpha_0 + \alpha_i + X'_{it,,Exp}\beta_{Exp} + X'_{it,TE}\,\beta_{TE} + X'_{it,RP}\,\beta_{RP} + \varepsilon_{it} \qquad (4)$$

where Q_i is the teaching quality rating observed on ratemyprofessors (We use no time subscript since our measure of quality is time-invariant.) As earlier, ϑ_i measures hotness, taking on a value of one if the number of hot ratings exceeds the number of nots. The parameter α_0 is a constant and the α_i represent university fixed effects. The vector $X'_{it,Exp}$ includes our measures of experience, rank and years since PhD. The controls for teaching effort, $X'_{it,TE}$, are being on sabbatical, being a department chair, university-level incentives to exert a high level of effort in teaching or publication such as merit pay, and a professor's number of ratings. $X'_{it,RP}$ is a measure of research productivity, and uses the same controls as our earnings regressions in Table 2. We estimate the hotness-quality effect separately for males and females so do not include a gender dummy.

Since teaching quality is time invariant, using all years of our panel adds little information on our dependent variable. We therefore restricted our sample to 2006 only. This also minimizes any potential problems arising from missing observations. As noted earlier, not every individual in our sample receives a rating on ratemyprofessors. In the earnings regressions, we were able to control for the presence of missing observations by including a dummy variable indicating that the individual had not been ranked on ratemyprofessors. For the teaching variables, we cannot use that strategy, as every individual with missing hotness information is also missing teaching quality information.[11] By restricting our sample to 2006 only, we create a more balanced panel, with fewer missing observations.

[11] In theory, we could correct for sample selection bias using the standard Heckman technique. Unfortunately, any variable plausibly correlated with not being included in ratemyprofessors (age, administrative duties, research output) would also be expect to affect teaching quality, leaving us without an instrument to carry out the Heckman correction.

Table 3 estimates equation 4, the impact of hotness on teaching quality, clarity, helpfulness and easiness for men and women.[12] We find, as do Bonds-Raacke and Raacke (2007) and Lawson and Stephenson (2005), a strong positive relationship between hotness and overall quality. The difference between our paper and others is that we also control for experience, effort and ability, confirming the robustness of the hotness/quality relationship.

Overall quality is an average of a professor's scores for helpfulness and clarity. Considering these two individually, we see that hotness has a stronger effect on men's clarity than their helpfulness, although the difference in coefficients is not statistically significant. For women hotness has a large, significant positive effect on perceived helpfulness. The effect of hotness on clarity, while positive, is not statistically significant. These findings are consistent with the idea that students are picking up on different attributes when they rate their male professors as hot as compared to when they rate their female professors as hot.

What about the effect of hotness on ease? Hotness has less of an effect on ease than clarity or helpfulness. As noted earlier, Johnson and Tassinary (2007) find that feminine women are more likely to be rated as attractive. Nelson (1995) argues that, in North American culture, rigor and toughness is associated with masculinity. Hence our result that the female professors voted not hot also received lower ease ratings – that is, were more rigorous and tough – is evidence in support of the connection between hotness and feminine behavior.

[12] We also estimated a cross section generalized least squares (GLS) regression using all years of data. The GLS results (available from the authors on request) are qualitatively similar but, due to the larger sample size, have much stronger statistical significance. The GLS point estimates for males are similar to the 2006 only estimates, but the point estimate of the effect of hotness on female professors' reported overall quality (0.52) and helpfulness (0.53) is smaller with the GLS results.

The average effect of hotness on perceived teaching quality is large– on a scale of 1 to 5, a hot professor is, on average, rated 0.67 points higher (overall results, not reported in Table 3). However being a good quality professor has few monetary returns. In the regression reported as case 7 in Table 1, the coefficients on ease, helpfulness and clarity are all positive but insignificant: easiness $2534 (standard error (s.e.) $1445), helpfulness $813 (s.e. $2043), clarity $814 (s.e. $1905). Moreover, comparing case 7 with the other cases, controls for teaching quality make little difference to the explanatory power of the earnings equation. So while the positive effect of hotness on teaching productivity is interesting, it does not explain the earnings premium associated with hotness.

6. Research productivity

We separately consider a number of elements of research productivity: number of publications per year, number of citations per year, receipt of a SSHRC grant, and number of co-authors (derived from publications per year). We assume each element of research productivity depends upon hotness (ϑ_i), experience ($X_{it,,Exp}$), teaching effort ($X_{it,TE}$), and other aspects of research achievement ($X_{it,R}^{j}$).

$$P_{i,t}^{j}=f(\vartheta_i, \alpha_0, \alpha_i, X_{it,,Exp}, X_{it,TE}, X_{it,R}^{j}, \varepsilon_{it}) \qquad (5)$$

where $P_{i,t}^{j}$ is the j^{th} element of research productivity, measured at time t for individual i. Our hotness, experience and teaching effort measures are the same as in the teaching regression. Our "other aspects of research" ($X_{it,R}^{j}$) capture the inherent relationships between the different measures of productivity. For example, we use the lagged quantity of SSHRC grant funding to predict the number of publications, citations and co-authors, on the grounds that research funding increases research output. These research measures change across research productivity measures to avoid simultaneity problems. For example, we include the number of past publications when we are explaining the number of citations (one can't be cited without publishing) but do not use citations to explain publications.

We have no censored observations for our research productivity measures. For the three quantitative measures (publications, citations, co-authors) we use a panel regression model that accounts for random effects to capture the remaining unexplained heterogeneity. The results of these regressions are shown in Table 4. We report estimates for males and females separately. Since the sample is predominantly male, the results for the entire sample are similar to the male results.

Columns 1 through 4 of Table 4 show the impact of hotness on the number of citations received each year and on log citations. The distribution of citations is highly skewed. A small number of people receive a very large number of citations, while most researchers' work is rarely cited. This makes log citations attractive as a dependent variable.[13] At the same time, since a relatively large number of individuals have citation counts of zero in any given year, logging citation counts causes us to lose a number of observations. Accordingly we repeated our analysis using citations (not logged).

Citations are often taken to indicate the quality of a publication (Hamermesh and Oster, 2002). However studies have documented gender differences in citation patterns, with men being more likely to be cited by men and women being more likely to be cited by women (Ferber, 1988). Since the majority of the economics profession is male, if men cite men more often, male economists will have higher citation rates overall, which is what is found by Ferber (1988) and, more recently, by Di Vaio, Waldenström and Weisdorf (2009). Table 4 shows that, for men, being hot has a minimal impact on the likelihood of being cited – the coefficient is positive but insignificant. For women, however, the effect of hotness on the log of citations is significant both economically and statistically – female hot professors have more than twice as many citations as the not. At the same time, however, since there is no

[13] We also experimented with a quantile regression approach in order to avoid losing information for the individuals who receive no citations in the social sciences citation index database in any one year. However the quantile regression approach did not converge for females.

significant impact of hotness on citations when we look at (not logged) citations, these results might be driven by a small number of individuals.

Unfortunately we did not collect data on the gender of citing authors so we do not know whether the hotness coefficient in column 2 of Table 4 reflects more citations by male or female economists. Moreover, it is possible that the link between hotness and citations is driven by some third factor, for example, working on an interesting research topic. But it is possible that, when a woman is hot, men find it easier to "recognize the merit" (Ferber, 1988) in her work. And one of the interesting things about this finding is the implication for future career success – once an article is cited, it begins to show up among the top hits in google scholar, which generates more citations and further recognition. It would be interesting to study this sample again in ten years time and find out whether being hot at an early stage in an academic career has implications for long-term success.

The next measure of productivity that we consider in Table 4 is the number of publications in the Econlit database. The database includes most Western English-language economics journal[14] publications and some books, but excludes, for example, book chapters or government reports. We consider annual publications, and estimate the effect of hotness on publications using generalized least squares with random effects.[15] For both men and women, once we control for experience, the effect of hotness on publication rates is positive but neither statistically nor economically significant – hot males publish one more article every seven years than the not, hot women one more article every 12.5 years. Translating this into financial terms, Sen et al (2009) find that the financial reward to an extra

[14] "The number of journals indexed in *EconLit* has grown from 182 periodicals in 1969 to over 750 journals today. Journals are selected for inclusion in *EconLit* on the basis of their economic content, which must be substantial or of equal emphasis in interdisciplinary journals." For list of journals included in Econlit, see http://www.aeaweb.org/econlit/journal_list.php.

[15] We also estimated the effect of hotness on obtaining a top 10 or top 21 journal publication. We found no significant effects.

publication ranges from $300 to $1,870 annually, depending upon the type of publication and the other controls included in the wage equation. Hence any indirect rewards to hotness via increased productivity are dwarfed by hotness's direct, financial return.

Columns 7 and 8 of Table 4 consider the impact of hotness on number of co-authors. Coauthorship rates matter because they have been hypothesized to underlie some of the gender disparities in the economics profession. Mcdowell and Kiholm Smith (1992) and, more recently, Boschini and Sjögren (2007) have found that economists tend to co-author with others of the same sex. Since there are relatively fewer women in the profession, this will tend to make it more difficult for women to find co-authors, which could, in turn, make it more difficult for women to publish.

The number of co-authors is derived from the publications data. Again, we use generalized least squares with random effects. There is some suggestion in the data that hot professors of either gender are more desirable co-authors. The results shown in Table 4 are not statistically significant, however Appendix Table C1 shows that, when we control for the potential selection bias, there is a weakly significant (p=0.06) positive effect of hotness on co-authorship for women.[16] Hamermesh and Oster (2002: 539) suggest that coauthorship has "consumption benefits," that is, people choose to coauthor with people whom they enjoy working with. Our results are consistent with the idea that the characteristics which make students rank professors as hot also make people desirable co-authors, but fall short of providing definite support for Hamermesh and Oster's hypothesis.

Our final test for the effect of hotness on productivity considered success in receiving a grant from the Social Sciences and Humanities Research Council (SSHRC), the primary Canadian funding council. The amount of grant funds received by successful applicants depends upon the individual's field

[16] For women, if we do not control for age, there is a statistically significant relationship between hotness and number of coauthors. However this may reflect a secular trend towards more co-authorship over time.

of research. Experimental economics, for example, is particularly expensive. Because we do not wish our measure of research funding success to be sensitive to the applicant's field, we define our dependent variable to be one during every year in which an individual held a SSHRC grant, and 0 otherwise. We use probit regression to estimate the probability of being in receipt of SSHRC funds. The controls employed are set out in equation 5, and are similar to those used in previous regressions.

Empirical results are reported in Table 5. We report separate estimates for males and females. The coefficient estimate on being hot is negative for both males and females, however when we control for selection into being rated hot, as described in Appendix C, the estimated impact of hotness is positive. In no case are the estimates statistically significant. Pooling the data together and running the regressions jointly for males and females did not change the significance of the 'hot' covariate. One way of understanding these results is to remember that the success rate for SSHRC grants is very low. Just 17 percent of men and 14 percent of the women in our sample receive any SSHRC funding. At the highly competitive level of SSHRC grant funding, hotness does not matter.

7. Conclusions

Male professors who are rated as hot by their students earn more, female professors do not. The results for men are primarily driven by mid-career and senior academics. The absence of a hotness premium for junior academics suggests that, conditional upon being hired for an academic position, appearance has a minimal impact on initial salaries. For established male academics, we hypothesize that hotness may be capturing masculine personality traits, such as assertiveness or confidence, that are associated with both sexual attractiveness and an ability to command higher salaries.

The results for research productivity are more mixed. There is some evidence that hot women are more likely to be cited, but this result is based on a small sub-sample. More generally, although the impact of hotness on measures of research productivity is usually positive, the relationship is not statistically significant. Both genders are much more likely to be rated as high quality teachers if they are hot, however the effect for women comes primarily from the positive effect of hotness on helpfulness, whereas hot men receive higher clarity scores.

We were genuinely surprised at the strength of the relationship that we found between hotness and earnings. We used numerous specifications and robustness tests to make sure that our results were not simply an artifact of a particular modeling choice. But every attempt to refine our empirical specification – introducing more productivity controls, controlling for the endogeneity of being rated – only increased the estimated returns to hotness. And *who* receives an earnings premium for hotness is even more interesting. For established male academics, hotness clearly pays, while women may face a conflict between being attractive and being financially well-rewarded.

Our results are not a pure measure of the effect of "beauty" on either academic salaries or professional productivity. Hotness measures some combination of physical attractiveness and other personality traits. It is possible that some people's hotness scores come more from their good looks, other people are hot because they are charismatic or likeable (and probably not bad looking). But whatever these attributes are that generate chili peppers on ratemyprofessors, they have real impacts on economic outcomes.

References:

Amemiya, Takeshi (1985), *Advanced Econometrics*, Oxford: Blackwell

Andreoni, James and Ragan Petrie (2008) "Beauty, gender and stereotypes: Evidence from laboratory experiments," *Journal of Economic Psychology* 29(1): 73-93, DOI: 10.1016/j.joep.2007.07.008

Barbezat, Debra A. (1987) "Salary Differentials by Sex in the Academic Labor Market," *Journal of Human Resources* 22(3): 422-28

--- (1991) "Updating estimates of male-female salary differentials in the academic labor market," *Economics Letters* 36(2): 191-95

Blackaby, David, Alison Booth and Jeff Frank (2005) "Outside Offers and the Gender Pay Gap: Empirical evidence from the UK academic labour market," *Economic Journal* 115(501): F81-F107

Bonds-Raacke, Jennifer and John D. Raacke (2007) "The Relationship between Physical Attractiveness of Professors and Students' Ratings of Professor Quality" *Journal of Psychiatry, Psychology and Mental Health* 1(2)

Boschini, Anne and Anna Sjögren (2007) "Is Team Formation Gender Neutral? Evidence from Coauthorship Patterns" *Journal of Labor Economics* 25(2): 325-365

Bowles, Hannah Riley, Linda Babcock, and Lei Lai (2007) "Social incentives for gender differences in the propensity to initiate negotiations: Sometimes it does hurt to ask," *Organizational Behavior and Human Decision Processes* 103(1): 84-103, ISSN 0749-5978, DOI: 10.1016/j.obhdp.2006.09.001

Chant, John (2005) "How we pay professors and why it matters" C.D. Howe Institute Commentary.

Di Vaio, Gianfranco, Daniel Waldenström and Jacob Louis Weisdorf (2009) "Citation Success: Evidence from Economic History Journal Publications" Univ. of Copenhagen Dept. of Economics Discussion Paper No. 10-01. Available at SSRN: http://ssrn.com/abstract=1531557

Felton, James, John Mitchell and Michael Stinson (2004) "Web-based student evaluations of professors: the relations between perceived quality, easiness and sexiness" *Assessment & Evaluation in Higher Education* 29(1): 91-108 DOI: 10.1080/0260293032000158180

Ferber, Marianne A. (1988) "Citations and Networking" *Gender and Society* 2(1): 82-89 http://www.jstor.org/stable/190470

Freng, Scott and David Webber (2009) "Turning Up the Heat on Online Teaching Evaluations: Does "Hotness" Matter?" *Teaching of Psychology* 36(3): 189 – 193 DOI: 10.1080/00986280902959739

Ginther, Donna K and Kahn, Shulamit (2004) "Women in Economics: Moving Up or Falling Off the Academic Career Ladder?" *Journal of Economic Perspectives* 18(3): 193-214

Gray, Mary and Barbara R Bergmann (2003) “Student teaching evaluations: Inaccurate, demeaning, misused” *Academe* 89(5)

Green,TG, FG Mixon JR, and Len Trevino (2005) “Have You Seen the New Econ Prof? Beauty, Teaching, and Occupational Choice,” in Franklin G Mixon (ed.) *Shaping the Learning Curve: Essays on Economic Education* New York: iUniverse Inc

Hamermesh, Daniel S. (2006) “Changing looks and changing ‘discrimination’: The beauty of economists,” *Economics Letters*, 93(3): 405-412 DOI: 10.1016/j.econlet.2006.06.004

Hamermesh, Daniel S. and Jeff E. Biddle (1994) “Beauty in the Labor Market,” *American Economic Review* 84(5): 1174-94

Hamermesh, Daniel and Sharon Oster (2002) “Tools or Toys: The impact of high technology on scholarly productivity,” *Economic Inquiry* 40(4): 539-555

Hamermesh, Daniel S. and Amy M. Parker (2005) “Beauty in the Classroom: Instructors’ Pulchritude and Putative Pedagogical Productivity,” *Economics of Education Review*, 24(4): 369-376

Hoopes, Jeffrey L. and Steve Albrecht (2007) “An Empirical Evaluation of Ratemyprofessor.Com” http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1019271

Johnson, Kerri L. and Louis G. Tassinary (2007) “Compatibility of basic social perceptions determines perceived attractiveness ,” *PNAS: Proceedings of the National Academy of Sciences* 104(12): 5246-5251

Kodrzycki, Yolanda and Pingkang Yu (2006) “New Approaches to Ranking Economics Journals,” *Contributions to Economic Analysis and Policy* 5(1), Art. 24

Kuhn, Peter and Kailing Shen (2009) “Employers’ Preferences for Gender, Age, Height and Beauty: Direct Evidence” NBER Working Paper No. 15564

Lawson, Robert A. and E. Frank Stephenson (2005) “Easiness, Attractiveness, and Faculty Evaluations: Evidence from RateMyProfessors.com,” *Atlantic Economic Journal* 33: 485–486

Martinello, Felice (2006) “University Revenues and Faculty Salaries in Ontario: 1970 and 2003” *Canadian Public Policy/Analyse de politiques* 32(4): 349-72

Mcdowell, John M. and Janet Kiholm Smith (1992) “The Effect of Gender-Sorting On Propensity To Coauthor: Implications For Academic Promotion” *Economic Inquiry* 30(1): 68-82

Miller, Geoffrey, Joshua M. Tybur, and Brent D. Jordan (2007) “Ovulatory cycle effects on tip earnings by lap dancers: economic evidence for human estrus?” Evolution and Human Behavior 28: 375–381

Mobius, Markus M. and Tanya S. Rosenblat (2006) “Why Beauty Matters,” *American Economic Review* 96(1): 222-235 doi: 10.1257/000282806776157515

Nyhus, Ellen K. and Empar Pons (2005) “The effects of personality on earnings,” *Journal of Economic Psychology*, 26(3): 363-384

Nelson, Julie A. (1995) “Feminism and Economics,” *Journal of Economic Perspectives* 9(2): 131-148 http://www.jstor.org/stable/2138170

Pagan, A. R. (1984), “Econometric issues in the analysis of regressions with generated regressors,” *International Economic Review* 25: 221-247

Riniolo, Todd C., Katherine C. Johnson, Tracy R. Sherman, Julie A. Misso (2006) “Hot or Not: Do Professors Perceived as Physically Attractive Receive Higher Student Evaluations?” *Journal of General Psychology* 133(1): 19 – 35 DOI: 10.3200/GENP.133.1.19-35

Sen, Anindya, Hideki Ariizumi, and Natasha De Sousa (2009) “How do we reward ourselves? Evidence on the returns to publications” paper presented at the Canadian Economics Association Meetings

Süssmuth, Bernd (2006) “Beauty in the classroom: are German students less blinded? Putative pedagogical productivity due to professors’ pulchritude: peculiar or pervasive?” *Applied Economics* 38(2):231-238

Timmerman, Thomas (2008) “On the Validity of RateMyProfessors.com,” *The Journal of Education for Business* 84(1): 55-61 DOI: 10.3200/JOEB.84.1.55-61

Warman,Casey, Christopher Worswick, and Frances Woolley (2010) “The Evolution of Male-Female Earnings Differentials in Canadian Universities,1970–2001” *Canadian Journal of Economics/Revue canadienne d’économique* 43(1): 347-372

Weeden, Jason and John Sabini (2005) “Physical Attractiveness and Health in Western Societies: A Review,” *Psychological Bulletin* 131(5): 635–653

Table 1: Impact of hotness on annual salary ($2006) – Censored (Tobit) Regressions							
	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7
Marginal effect of hotness dummy: full support	6,860	7,731	12,902[c]	12,910[c]	12,957[c]	9,219[b]	7,508[a]
Standard error	6582	6444	4561	4564	4210	3860	4191
Marginal effect of hotness dummy conditional on being uncensored	1947	2193	3239[c]	3241[c]	3153[c]	2232[c]	1785[a]
Standard error	1791	1743	1009	1009	900	856	928
Marginal effect of hotness dummy on probability of being uncensored	0.062	0.071	0.167[c]	0.167[c]	0.066[c]	0.113[c]	0.091[a]
Standard error	0.058	0.057	0.051	0.051	0.046	0.042	0.047
Dummy if individual has no rating on www.ratemyprofessor.com		*	*	*	*	*	*
Gender		*	*	*	*	*	*
University and year dummies			*	*	*	*	*
Lagged: journal publications, journal publications squared, citations, SSHRC funding. Book. Dummies for: Department chair, U.S. PhD, other non-Canadian PhD. Number of ratings.			*	*	*	*	*
Years since PhD, Years since PhD squared			*	*	*		
Dummies if university is unionized, has salary cap, or merit pay				*	*	*	*
Dummies for assistant, full professor, other rank, rank missing					*	*	*
clarity, helpfulness and ease of professor (1 to 5 scale)							*
Log likelihood	-14,184	-14,146	-13,186	-13,186	-13,157	-13,246	-13,240

Notes: [a] denotes significance at p=0.10, [b] denotes significance at p=0.05, and [c] significance at p=0.01. The data is assumed to follow the observability rule in equation (2) above. Standard errors are clustered at the individual level. All of the above regressions have been run using a sample of 3,455 observations based on a panel of 493 individuals. The estimation method is a generalization of the Tobit model introduced by Tobin (1958) – see Wooldridge (2002, Chapter 16).

Table 2: Effect of hotness (marginal effect) on annual salary of Ontario economics professors, 2006 Cdn $, selected sub-groups – censored regressions.						
	Male	**Female**	**Year PhD >2000**	**Year PhD >1995**	**Year PhD >1985**	**Year PhD >1975**
Marginal effect of hotness dummy: full support	13,725[c]	-1,524	-2,396[c]	272	13,140[c]	13,323[c]
Clustered standard error	4,677	6,171	913	798	909	4,390
Marginal effect of hotness dummy conditional on being uncensored	3,638[c]	-89	-67[c]	5	967[c]	2,384[c]
Clustered standard error	1,078	365	26	16	60	681
Marginal effect of hotness dummy on probability of being uncensored	0.189[c]	-1.10×10^{-4}	-3.02×10^{-8}[b]	3.49×10^{-12}	0.00406[c]	0.125[c]
Clustered standard error	0.056	5.02×10^{-4}	1.49×10^{-8}	9.89×10^{-12}	0.00013	0.030
N (observations)	3032	423	229	547	1503	2399
N (individuals)	416	77	86	150	272	364
Log likelihood	-12,309	-798	-385	-1,100	-3297	-7078

Notes: [a] denotes significance at p=0.10, [b] denotes significance at p=0.05, and [c] is significance at p=0.01. The data is assumed to follow the observability rule in equation (2) above. Standard errors are clustered at the individual level. The estimation method is a generalization of the Tobit model introduced by Tobin (1958) – see Wooldridge (2002, Chapter 16). The controls used correspond to "Case 5" of Table. Controls not reported include: years since PhD; years since PhD squared; number of ratings; lagged number of journal publications and publications squared; number of books published, lagged SSHRC funding; lagged citation count; dummy variables for gender (columns 3 through 6), academic rank, year, university, union, salary cap, merit pay schemes, and a dummy indicating individual was not rated on www.ratemyprofessors.com.

Table 3: Impact of a “hot” appearance (marginal effect) on students’ evaluations of teaching performance, 2006 only								
	Quality		Clarity		Helpfulness		Easiness	
	Male	Female	Male	Female	Male	Female	Male	Female
Hot	0.72[c]	0.81[b]	0.76[c]	0.55	0.68[c]	1.02[c]	0.12	0.43[a]
Standard error	0.16	0.30	0.17	0.36	0.16	0.28	0.13	0.23
Adjusted r^2	0.11	0.40	0.10	0.17	0.11	0.54	0.05	0.44
N	229	46	229	46	229	46	229	46

Notes: [a] denotes significance at p=0.10, [b] denotes significance at p=0.05, and [c] is significance at p=0.01. The evaluations of teaching, quality, clarity, helpfulness, and easiness, range from 1 to 5. Controls not reported include: academic rank dummies; dummies for sabbatical leave, whether the individual is a department chair and whether the individual has a PhD from a U.S. or any other non-Canadian university; university dummies; dummies for union, salary cap, merit pay schemes; lagged publications and publications squared; lagged SSHRC funding; lagged citation count; years since PhD and its square; number of ratings on ratemyprofessors. Since we have no information on the teaching performance of those not rated on ratemyprofessors, we do not include a “not rated” dummy.

Table 4: Impact of hotness on productivity measures								
Dependent Variable	Log citations		Citations		Annual publications		Coauthors	
	Male	Female	Male	Female	Male	Female	Male	Female
Hot	0.19	1.86[c]	0.08	-0.09	0.14	0.08	0.14	.19
Robust standard error	0.15	0.44	0.82	1.41	0.13	0.14	0.15	0.17
Rho	0.13	0	0	0	0.21	0.07	0.24	0
r^2: within	0.41	0.70	0.10	0.41	0.02	0.10	0.01	0.05
r^2: overall	0.39	0.82	0.14	0.47	0.09	0.24	0.09	0.22
r^2: between	0.35	0.92	0.19	0.62	0.17	0.29	0.17	0.35

Notes: [a] significant at p=0.10, [b] significant at p=0.05, and [c] is significant at p=0.01. The dependent variables in the above regressions are log(citations), citations, annual publications, and the number of co-authors. All zero values are dropped when the dependent variable is log(citations). Otherwise all regressions are based on the entire sample. They were carried out using Generalized Least Squares with random effects. Robust standard errors are reported. Controls not reported include: lagged SSHRC funding; a dummy if the individual was not rated on www.ratemyprofessors.com; academic rank dummies; dummies for sabbatical leave, whether the individual is a department chair and whether the individual obtained a PhD from a U.S. or any other non-Canadian university; year dummies; university dummies; dummies for union, salary cap, merit pay schemes; years since PhD and its square term; number of ratings on ratemyprofessors; publication of a book. Lagged publications and publications squared are included in the citation but not the publications and coauthors regressions.

Table 5: Impact of hotness on probability of receiving a SSHRC grant: Probit analysis				
	Males		Females	
	observed selection into not being rated	Predicted selection into not being rated	observed selection into not being rated	Predicted selection into not being rated
Hot (probit coefficient)	-0.006	0.62	-0.147	0.113
Clustered standard error	0.118	0.116	0.383	0.395
Hot (marginal effect)	-0.001	0.010	-0.027	0.023
Clustered standard error	0.020	0.021	0.182	0.159
N	2529	2529	267	267
LR Chi^2	833	838	99	96

Notes: [a] denotes significance at p=0.10, [b] denotes significance at p=0.05, and [c] is significance at p=0.01. The dependent variable in the above regressions is 1- if the individual has a SSHRC grant, 0 – otherwise. The above table reports actual probit coefficient estimates as well as the marginal effects of being Hot. Standard errors are beneath marginal effects estimates and clustered at the individual level. "Observed selection" regressions include a dummy if the individual was not rated on www.ratemyprofessors.com; "predicted selection" regressions Controls not reported include: academic rank dummies; dummies for sabbatical leave, whether the individual is a department chair and whether the individual is from a U.S. or any other non-Canadian university; year dummies; university dummies; dummies for union, salary cap, merit pay schemes; and lagged publications and publications squared.

Table A1: Descriptive Statistics. Notes: includes all years. In this table, professors receiving no ratings are coded as zero for hotness, ease, clarity and helpfulness.								
ENTIRE SAMPLE (N=3455)			HOT ONLY (N=349)		MALES (N=3032)		FEMALES (N=423)	
Variable	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Dependent variables								
Salary ($2006) (≥$100,000 only)	$127,017	$19,550	$129,983	$265,412	$127,672	$19,783	$117,418	$12,313
Proportion earning <$100,000	0.69	0.46	0.64	0.48	0.67	0.47	0.84	0.37
Clarity	2.44	1.66	3.95	0.69	2.45	1.68	2.34	1.45
Helpfulness	2.47	1.66	3.95	0.66	2.48	1.69	2.39	1.46
Easiness	2.13	1.40	3.04	0.56	2.13	1.42	2.12	1.27
Annual citations	1.26	5.53	1.91	6.23	1.35	5.83	0.56	2.36
Annual publications	0.52	0.92	0.69	1.08	0.53	0.95	0.41	0.72
Coauthors	0.49	1.18	0.74	1.39	0.51	1.22	0.34	0.84
Any SSHRC funding	0.17	0.37	0.27	0.44	0.17	0.38	0.14	0.35
Explanatory variables								
Hot	0.10	0.30	1.00	0.00	0.09	0.29	0.16	0.37
Not rated	0.27	0.44	0.00	0.00	0.27	0.44	0.23	0.42
Yrs since PhD	17.87	10.52	11.69	9.93	18.86	10.44	10.80	8.02
Yrs since PhD^2	430.01	406.84	235.02	346.08	464.79	412.68	180.69	245.99
Below Asst.	0.01	0.09	0.01	0.11	0.01	0.08	0.02	0.15
Assistant	0.21	0.41	0.46	0.50	0.18	0.39	0.40	0.49
Associate	0.29	0.45	0.25	0.43	0.28	0.45	0.38	0.48
Full	0.46	0.50	0.28	0.45	0.50	0.50	0.18	0.38
Rank missing	0.03	0.16	0.00	0.00	0.03	0.17	0.02	0.14
Sabbatical	0.11	0.32	0.12	0.33	0.12	0.32	0.09	0.28
Publications squared	1.11	3.76	1.63	4.07	1.17	3.95	0.69	1.84
Chair	0.05	0.23	0.03	0.16	0.06	0.23	0.05	0.21
Num. Ratings	16.49	22.09	15.61	13.69	16.96	23.02	13.11	13.31
Book	0.02	0.12	0.01	0.08	0.02	0.13	0.00	0.07
Male	0.88	0.33	0.81	0.40	1.00	0.00	0.00	0.00
SSHRC funding ($'000s)	4.95	24.30	7.10	21.51	5.18	25.57	3.27	11.63
US PhD	0.46	0.50	0.35	0.48	0.47	0.50	0.37	0.48
Other PhD	0.08	0.28	0.06	0.24	0.08	0.27	0.12	0.33
Union	0.60	0.49	0.54	0.50	0.58	0.49	0.68	0.47
Salary cap	0.33	0.47	0.16	0.37	0.31	0.46	0.43	0.50
Merit pay	0.86	0.35	0.85	0.35	0.87	0.34	0.79	0.41

Appendix Table A2: Robustness checks									
	Tobit regression			Tobit regressions, “Not rated” dummy			Probit	GLS	GLS
	Without selection control	“Not rated” dummy	“Not rated” prediction	2006 only	Nominal salaries	1996$, lower limit $100,000	Over $100K=1	Salaries over $100K only	Log salary, Over $100K only
Hot: marginal effect (full support)	13,442[c]	12,957[c]	13,899[c]	6,727[a]	10,274[c]	9,753[b]	0.15[c]	6535[c]	0.045[c]
Clustered standard error	4,154	4,209	4,205	3,796	3,809	4,237	0.03	2517	0.017
Hot: marginal effect (uncensored only)	3,288[c]	3,154[c]	3,453[c]	3,089[a]	2,323[c]	1,515[c]	n/a	n/a	n/a
Clustered standard error	888[c]	900[c]	906[c]	1622	757	578	n/a	n/a	n/a
Hot: marginal effect on probability of being censored	0.173	0.165	0.184	0.132[a]	0.153[c]	0.072[c]	n/a	n/a	n/a
Clustered standard error	0.045	0.046	0.047	0.078	0.047	0.023	n/a	n/a	n/a
N (observations)	3455	3455	3455	328	3455	3455	3424	1080	1080
N (individuals)	493	493	493	328	493	493		271	271
log likelihood	-13,158	-13,157	-13,135	-2,210	-12,821	-9,478	-1268	X^2=879	X^2=1157

a) Significant at p=0.10, b) significant at p=0.05, c) significant at p=0.01
All specifications correspond to “Case 5” of Table 1. Controls not reported: gender, dummy indicating no rmp rating, rank, year dummies (10), university dummies (15), union, salary cap, merit, number of ratings, lagged publications and publications squared, lagged SSHRC funding, lagged citation count, book, years since PhD and years since PhD squared.

Appendix A:

Easiness:

Hard Easy

Helpfulness:

Useless Extremely Helpful

Clarity:

Incomprehensible Crystal

Interest level prior to attending class:

None at all It's my world!

Textbook Use:

Low High

Textbook Used*: (ex. ISBN#: 079074272X)

Grade: Please Select

Attendance:

Mandatory Not Mandatory

Prof Status:

Still Teaching Retired/Gone

Appearance: (just for fun)

Hot Not

Class: (i.e. HIST 101, ACCT 202) CLAS SECT

Comments:
Please keep comments clean. Libelous comments will be deleted. Guidelines

Characters Typed: 0 (Limit 350 per rating)
Remember, **YOU ARE RESPONSIBLE** for what you write here.

Rating Categories

RateMyProfessors.com's ratings categories:

Easiness - Some students may factor in the easiness or difficulty of the professor or course material when selecting a class to take. Is this class an easy A? How much work do you need to do in order to get a good grade? *Please note this category is NOT included in the "Overall Quality" rating.*

Helpfulness - Helpfulness is defined as a professor's helpfulness and approachability. Is this professor approachable, nice and easy to communicate with? How accessible is the professor and is he/she available during office hours or after class for additional help?

Clarity - A professor's organization and time management skills can make a great difference on what you get out of the class. How well does the professor teach the course material? Were you able to understand the class topics based on the professor's teaching methods and style?

Overall Quality - The Overall Quality rating is determined by the average rating of the Helpfulness and Clarity given by all users. An overall rating of 3.5 to 5 is considered good (yellow smiley face). An overall rating of 2.5 to 5 is considered average (green smiley face). An overall rating of 1 to 2.5 is considered poor (blue sad face). The Easiness rating is NOT included when calculating the Overall Quality rating.

Rater Interest - There is always that one class everyone recommends taking before graduating. As a student, how interested were you in the class, BEFORE taking it? Or how interested were you in taking this course from this specific professor.

Source: http://www.ratemyprofessors.com/categories.jsp

Appendix B: Data Sources

Data Source	Date accessed	Variables
Department web sites, updated with information from university calendars and Canadian Economics Association newsletter	Summer 2004 – 2008	Names of individuals to include in sample, identification of individuals holding administrative positions, academic rank.
Department web sites, personal communication	Summer 2004 – 2008	Department Chair ("Chair"). Note: administrative not research chairs. Administrators at levels above Chair are excluded. These are typically individuals with University level administrative responsibilities
PROQUEST Dissertations and Theses database	Summer 2004 – 2008	Years since PhD obtained, PhD from US institutions, PhD from non-US/Canadian institutions
Departmental Home Pages, gender references in www.ratemyprofessor.com	Summer 2004 – 2008	Male
How we pay professors and why it matters C.D. Howe Institute Commentary, Nov, 2005 (John Chant)	Summer 2008	Dummies on university level compensation schemes – whether the institution is unionized, implements salary caps, and/or merit pay
ECONLIT	Summer 2004 – 2008	# publications in peer reviewed journals per year, # publications in peer reviewed journals prior to 1996, number co-authors worked with, number of books published. The number of co-authors is the sum of all co-authors listed in all journal articles published in each year.
"Winning Research" search at www.sshrc.gc.ca	Summer 2009	SSHRC
Social Sciences Citation Index	Summer 2008	Individual citations (total number of citations received by an individual on an annual basis for cumulative research)
Ontario Ministry of Finance website, accessible at http://www.fin.gov.on.ca/en/publications/salarydisclosure/2008/	Summer 2004 – 2008	Annual (calendar year) salary
www.ratemyprofessors.com	Summer 2009	Dummy variables for hotness, helpfulness, clarity, and ease

Appendix C: Selection issues

Not all faculty members are rated on http://www.ratemyprofessors.com/. If the selection of professors for ratings is non-random, it may have a systematic impact on the reported results.

We use two alternative measures of selection into not being rated to control for this problem. First, we use an observed variable (the dummy for unrated) that identifies the nonrated professors and proxies this selection into rating. Second, we used a generated variable that is a prediction of the true probability of being not rated, obtained by estimating a selection equation that has variables that are also found in the original model, but also some exclusion restrictions, or variables that are important for the selection onto rating but may not be important for the other models that we test. Accordingly, we estimate the probability of being nonrated by using the same exclusion restrictions (editor, administrative dummy, publications prior to 1996) for each model that we test. In the case of the generated regressor, or predicted probability of not being rated, the identification may not require strong exclusion restrictions as the model can be also weakly identified through the functional form of the probability of being nonrated which is a nonlinear function of the regressors used. Note also that the identification requirements (strictly exogenous exclusion restriction) for a linear predictor are more stringent than the requirements for the non-linear predictor used here.

Comparison of results obtained using the observed variable (not rated) with those obtained using the generated regressor (predicted probability of being rated) raises two significant empirical issues. The predicted probability is measured with error as it is generated using a selection equation. Therefore, its use adds measurement variability to the estimating equation which has impacts for the parameter estimates of the overall model that we choose to test. Pagan (1984) in his analysis of regressions containing generated regressors, shows that if the generated regressor is consistent, the parameter estimate of the generated regressor will also be consistent, as long as the original model is correctly specified. To deal with this issue, we use the model that better explains the changes in salaries before introducing the predicted probability of not being rated. If the generated regressor estimates the subjective probability consistently, there is no inconsistency in the parameter estimates. At worst the standard errors of both the generated regressor and the non-generated regressors in the regression equation will be increased by the generated regressor's inconsistent standard errors. Therefore, if we observe non-significant changes in the standard errors of the non-generated regressors due to the introduction of the generated- regressor, the noise induced by the error in variable associated with the generated regressor will have a minimum impact on the other variables.

On the other hand, proxing the selection of rating with the nonrated dummy uses an observed variable that can be measured without error. Moreover, the nonrated dummy may not be free of endogeneity problems for all the models that we are testing. Therefore, for the models for which we observe a lack of improvement in fit when the dummy is included, we may have simultaneity or other types of endogenety problems.

In the main body of this paper we have, with the exception of Table 5, reported results controlling for selection into rating using the not rated dummy, an observed variable that takes on a value of one is the professor has not ratings in ratemyprofessors. Table C1 reproduces the results of Tables 1, 2 and 4 using the generated variable, rather than the observed variable, to correct for selection into ratemyprofessors. In most cases, the effect of hotness is amplified somewhat, but there are few changes in the sign or significance of the results. For the results in Tables 1 and 2, there are minimal changes to the log likelihood function when using the predicted value of not being rated, as opposed to the actual value. For this reason, and to make the results more transparent and easier to interpret, we use the observed rather than the predicted value of not being rated to control for selection.

Table C1: Key results using predicted value of "hot rated"

		hot	**s.e.**	**N**	**Log likelihood**
Table 1, Case 7	All	13,889[c]	4,205	3455	-13,135
Table 2					
	Males	15,116[c]	4,669	3032	-12,287
	Females	-4,887	7,081	423	-796
	PhD>2000	-1,043	834	229	-386
	PhD>1995	1,320	802	547	-1,100
	PhD>1985	13,184[c]	862	1503	-3,297
	PhD>1975	13,437[c]	4,429	2399	-7,072
Table 4					**Chi^2**
Log citations	Male	0.165	0.148	581	760
	Female	1.53[c]	0.309	61	.
Citations	Male	-0.150	0.813	1003	.
	Female	-0.228	1.47	128	.
Annual publications	Male	0.186	0.114	3032	185
	Female	0.163	0.144	423	.
Coauthors	Male	0.175	0.142	3032	154
	Female	0.305[a]	0.165	423	.

a) Significant at p=0.10, b) significant at p=0.05, c) significant at p=0.01. For full listing of other covariates, see Tables 1, 2 and 4 in main body of text.