Observable persuaders: A longitudinal study on the effects of quality signals in the contemporary visual art market

Kackovic, M.

Publication date
2016

Document Version
Final published version

CHAPTER 2

PICKING THE CREAM OF THE CROP:

A Study on the Effectiveness of Jury Selection to a Prestigious Art Academy

Kackovic, M., Hartog, J., Van Ophem, H., Wijnberg, N.M.

An earlier version of this chapter was presented at the 2016 Cultural Economics Conference in Valladolid, Spain as a conference paper presentation.

ABSTRACT

Research on the selection of applicants to academic programs has generally focused on formal admission criteria, emphasizing the role of objective, unequivocal measurements of scholastic merit. We study a new domain: selection to an elite visual arts program, where selection procedures are semi-formal, with specific focus on assessing applicants' latent or hard-to-observe quality and developmental potential, rather than simply past achievements. We analyze 11 years of comprehensive selection committee deliberation notes, and identify multiple selection criteria used by committee members to select potentially excellent individuals from a large pool of applicants. Based on applicants' scores on these criteria and their biographical data, we estimate (1) acceptance to the program and (2) future performance, i.e., auction sales and reputational rankings of artistic standing using switching regressions. We find that multiple selection criteria have predictive value: receiving high scores on criteria, as well as the final jury assessment, are positively correlated with admission and better reputational rankings, but not with sales.

Keywords:
selection, artistic labor market, prestigious education, switching regressions

2.1 INTRODUCTION

Selection processes focused on choosing the *right* applicant from a pool of applicants are common in many settings, ranging from hiring employees or managers to choosing participants for scholarly or professional programs. In these selection processes decisions have to be taken under conditions of uncertainty. Information about the present characteristics of the applicants is always imperfect, and even more so about their future potential in an uncertain future environment. To reduce uncertainty, many selection committees rely on selection criteria to help them choose potentially excellent applicants among the many competing alternatives. Choosing the right applicant is important because the costs associated with making the *wrong* selection can have far reaching negative financial and/or reputational impact on the selectors, applicants and the competitive landscape of a market.

Throughout history there have been striking examples of selection misjudgments. For instance, at age 16 Albert Einstein was first rejected by the prestigious Swiss Federal Polytechnic Institute (ETH) in Zurich only to be accepted a year later; Fritz Kreisler, acclaimed as one of the world's top violinists of the twentieth century was rejected by the Wiener Philharmoniker and initially reacted by studying medicine instead; two time Academy Award winner, Steven Spielberg, was rejected twice from the elite film and theater school at the University of Southern California. These types of misjudgments are not rare. To reduce the risk of rejecting potentially excellent applicants or selecting inferior ones, selection committees often assess the information provided during the selection process along different dimensions, using multiple criteria of quality. Designing a selection procedure, developing such criteria, and involving qualified selection committee members is both a time-consuming and a costly process (Klitgaard, 1985), but especially at elite institutions, choosing a limited number of potentially excellent candidates from large pools of applicants might just be worth spending resources on.

The core aim of this paper is to gain a better understanding of the predictive value of multiple selection criteria used by a selection commission to evaluate applicants' latent or hard-to-observe quality and developmental potential. We identify these criteria and analyze the extent to which they predict: 1) admission to the program, 2) later sales performance and reputational rankings for both the accepted and rejected applicants, and 3) the extent to which these criteria overlap in predicting both selection and future performance. We analyze the final round of a selection process, in a procedure where the large number of applicants necessitates a mostly administrative pre-selection.

Extant research on selection processes has focused on settings with formal admission procedures, in which selection decisions are based either on objective measures, e.g., grades and standardized test score, (e.g., Grove and Wu, 2007; Ehrenberg and Marvos, 1995; Van Ours and Ridder, 2003; Klitgaard, 1985), or subjective measures, but rooted in formal structures and binding procedures on the basis of expert opinions (e.g., Glejser and Heyndels, 2001; Ginsberg and Van Ours, 2003). We focus on a setting in which the selection process is semi-formal, and admission decisions are based on applicants' latent quality and potential future development, making this an excellent setting to study the predictive value of multiple selection criteria.
The empirical setting of our study is the contemporary visual arts market, which is known to

have an oversupply of producers, uncertain demand and no objective measures of determining product or producer quality (Caves, 2000; DiMaggio, 1987; Yogev, 2010; Velthuis, 2013). We study the selection process at the *Rijksakademie van Beeldende Kunsten* (RABK), an internationally renowned and prestigious visual arts residency located in Amsterdam, the Netherlands. The goal of RABK is to select the most talented visual artists and offer them a platform where research, experimentation, innovation and critical discourse are central, complemented with technical facilities, workshops, podia for artistic presentations and networking opportunities (Rijksakademie Annual Report, 2014). To fulfill this goal RABK invests heavily in talent scouting and development, which starts with a rigorous selection process. During our 11-year observation period, 9249 visual artists from across the globe applied to RABK and on average 27 were accepted every year. In accordance with RABK goals, the selection committee, consisting of internationally renowned visual artists, spend five weeks every year selecting a few potentially talented visual artists to allow them to further develop and deepen their art practice (www.rijksakademie.nl). Being admitted to an elite program, such as RABK, can positively affect later performance. To address potential selectivity in our empirical analysis, we use switching regressions based on parametric and semi-parametric two-step selection models (Heckman, 1974,1976; Newey, 1999).

The paper is organized as follows: In Section 2 we discuss existing literature on selection processes to academic and other specialized programs. Section 3 provides a detailed description of the empirical setting and evolution of selection procedures at RABK. Section 4 presents the data and descriptive statistics. Section 5 discusses differences among performance of accepted and rejected artists. Section 6 analyses the admission decision in greater detail, and presents the switching regressions as our analytical strategy to test if admitted candidates perform better than the rejected ones. Section 7 presents results from robustness tests and simulates alternative admission procedures and their consequences. We present our conclusions in section 8.

2.2 THEORETICAL FRAMEWORK

Predicting success in economic and artistic life is notoriously difficult, as several strands of literature indicate. Occupational psychology claims that the benefits of testing job applicants, for higher productivity, can outweigh the cost of testing (Cascio, 1999) even though explained variance in job performance is not high. Van der Maesen de Sombreff (1992, p 11) lists results from meta analyses of several types of variables. The highest score is for cognitive abilities, biographical data and structured assessment of training and experience, with each an average R^2 of about 0.25. Structured interviews score slightly lower, at about 0.16. Unstructured interviews score no higher than 0.06, personality questionnaires; unstructured assessments of training and experience and average school grades score even lower. Wright, Lichtenfels, Pursell (1989) also note that structured interviews give better results than unstructured interviews. In another stream of literature, i.e., labor economics, success is measured as earnings. Individual earnings functions usually only explain a tiny fraction of the variance (Hartog, 2001; Taubman, 1975). The most powerful variable is (years of) schooling. Abilities and personality variables have an impact but their explanatory power is weak (Hartog, 2001); although with sufficiently strong econometric modeling, personality variables can make a significant contribution (Borghans, Duckworth, Heckman, Ter Weel, 2008).

The literature on predicting academic success, at the advanced level, generally finds that grades or standardized test scores affect selection decisions. Although the empirical literature is quite modest about the benefits of an elite education, a few papers consider the success of selective admission to economics PhD programs. Ehrenberg and Mavros (1995) find that GRE scores predict selection, but fail to predict PhD completion or time-to-degree for economic PhD students at Cornell University. Groove and Wu (2007) analyze the probability of PhD completion and research productivity of all applicants to a top-five economics PhD program in the US in 1989. They find that the committees' subjective ratings of applicants have clear predictive value, 'but do not encompass the information contained in the applicants' file.' Research on committees comprised of peers has shown that peer based selection is a useful way of maintaining high quality standards (Kostoff, 1997) because peers are best able to judge novelty and contribution of output (Eisenhart, 2002). Past studies have shown that agreement between peers involved in selection processes is strongly correlated with future positive selection outcomes (Cole and Simon, 1981; Bornmann and Daniel, 2005). Nevertheless, these studies relied heavily on objective criteria such as standardized test scores and high school grades (Grove and Wu, 2007; Ehrenberg and Marvos, 1995; Van Ours and Ridder, 2003), the status of the applicants' previous educational institution, and letters of recommendation (Klitgaard, 1985; Grove and Wu, 2007); there is also specific focus on the effects of financial support (Ehrenberg and Mavros, 1992; Booth and Satchel, 1995) and supervision (Van Ours and Ridder, 2003; Hilmer and Hilmer, 2012; De Valero, 2001). Grove and Wu, (2007) have noted that the highest predictive value obtains from subjective rating combined with administrative data, i.e., committee ratings alone predict less accurately than all administrative variables combined. In short, these studies found that past scholastic output affects future performance, e.g., job placement, but fail to offer any reasons for this strong correlation. To try to gain a better understanding of this relation, Athey, Katz, Krueger, Levitt and Poterba (2007) ask the following questions: Does education teach useful skills that are employable later in life? Or are students with higher grades more self-confident, thus more likely to find a better job? Or are cognitive abilities coupled with diligence and creativity truly important traits to achieving future success? Yet, they too are unable to unambiguously explain this correlation, and instead conclude, "our results suggest that there is not an easy recognizable star profile or single path to success" (Athey, et al., 2007, p. 518).

In a different stream of literature focusing on predicting success at contests and competitions with highly structured selection processes, Glejser and Heyndels (2001) found that distortions of efficiency such as strategic voting by jury members, self-selection of participants, discrimination (by gender), type of work performed, and conditions of performance, e.g., time of day and day of the week can affect the outcomes of candidates competing in the Queen Elizabeth (QE) Contest. For instance, Ginsberg and Van Ours (2003) found that candidates who performed later during this contest had higher rankings, regardless of their quality[13]. In a related study on elections of Econometric Society Fellows, Hamermesh and Schmidt (2003) conclude that highly structured selection processes are often influenced by other factors that may not be directly related to the candidate's research quality, yet affect performance outcomes.

This review of related literature indicates that we should not anticipate a very precise and effective selection procedure, in particular as the RABK selection process is fairly informal, as a non-structured discussion based on subjective assessments: "the final selection is made collectively based on each group's assessment". There is no formal voting, no explicit grading on explicit criteria. This is in contrast to objective formalized decision making, with thresholds for grades on objective, unequivocally measured variables (diploma yes/no, school grades, publication scores), or subjective formalized decision making with binding procedures on basis of expert opinions, as in the QE Contest. But as these features are not unique to the RABK it is certainly worthwhile to analyze the workings and results of the process.

2.3 EMPIRICAL SETTING

2.3.1 Evolution of the selection process

RABK was established in 1870 as a classical art academy, focusing exclusively on teaching skills used to make art. In the late 1980s, RABK began transitioning from a classical art academy to becoming a two-year artists' residency program, in which the focus from teaching the technicalities of fine art disciplines, such as painting and sculpture, shifted to a new domain: research, experimentation, innovation and critical discourse (Rijksakademie Annual Report, 2014).

Since their development as an art residency program, there has been a steady evolution in their selection process. There are three distinct phases in which selection procedures changed. In the first phase, which lasted from 1990-1993, a jury member specialized in a specific art genre solely evaluated applicants within that core discipline. In 1992 no selection took place because of organizational restructuring and location change. From 1993 onward, the residency period of two years was based on the calendar year, January to December, and no longer an academic year.

In the second phase, which lasted from 1994-1998, there were two jury groups, namely: one specialized in two-dimensional work such as painting, drawing, photography and graphic design, and the other one specialized in three-dimensional work such as sculpture but also film and art video.

13 Chan, Frey, Gallus and Torgler (2103) give good references to the literature on contests and competitions.

Each jury group first considered applicants from their core disciplines, then jury members from each respective group considered the other groups' applicants, and lastly, based upon discussions among all jury members, applicants were selected for an interview. At the interview phase, the jury of the core discipline of the applicant made the final selection. In 1998, an extra Dutch selection was added so that Dutch applicants could bring in original work for viewing.

And in the third phase, starting in 1999 until 2010 the last date of our data collection, selection was not based upon core disciplines but rather the jury consisted of multi-disciplinary groups who assessed the work from all applicants. A pre-selection round was also introduced. In 2006 the applicants in the selection procedure were sub-categorized into Dutch and foreign. And finally, in 2009, due to the large amount of applicants a (pre)-pre selection round was introduced which included applicants who are not in the target group. Our observation period begins in 1995 and ends in 2008; we do not have observations in the first phase, and we control for two years from the second phase by creating a dummy variable coded one for the years 1995 and 1997, and otherwise zero; the rest of our observations are in the third phase.

2.3.2 Selection Rounds

Since the 1990s, RABK has had an interview round in their selection process, regardless of the number and types of preceding rounds, although as the number of applicants increased, so did the number of selection rounds. The number of applicants tripled from 1995 to 2010, necessitating the introduction of a (pre-) pre selection round and a pre selection round for both Dutch and foreign applicants preceding the selection and interview rounds.

Particularly for international applicants, the (pre)-pre selection round consists of individuals who are not in RABK's target group. These applicants are younger than 26 or older than 36 years old or have less than 2 years art practice experience or do not have a letter of recommendation. This round is characterized by quick viewing of (audio-) visual work and usually most applicants in this round will be rejected. The ones who do proceed to the next round usually have either been to an interview in a previous year or are younger artists whose visual work is striking.

Those applicants who do fit the explicit age and experience criteria are included in the so-called pre-selection. This round is sub-divided into Dutch and foreign candidates, and is composed of individuals who are in the target age group of 27 to 36 years of age and have two or more years of individual art practice experience but do not have a letter of recommendation. Both the (pre)-pre selection and the pre-selection rounds are characterized by very quick viewing of artworks projected onto a screen. Main selection criteria in these two rounds are technical skill and autonomous style, i.e., not heavily influenced by a particular style, school or tutor. Artists with letters of recommendations are exempt from either of these rounds. Interestingly, artists who are affiliated with a notable gallery or have an established art practice and display their work at prestigious international exhibitions are usually rejected a this phase of selection. This is a very conscious decision made by the selection committee because of these artists' advanced career trajectories and lack of fit with the objectives of RABK of identifying possibly talented individuals and helping them develop this potential.

Artists then progress to the selection round, which consist of two-parts: first, the selection committee members view submitted artworks per applicant, either digitally by being projected

onto a screen or original works. At the end of the first part of this round, jury members are explicitly asked their opinion about applicants' artworks and each juror makes a short list of favorite applicants[14]. In the second part, applicants' description about their work and their expectations from RABK are read out loud, and letter(s) of recommendation are discussed in detail. Recommendation letters are considered very carefully, and if, for instance, an applicant with a letter of recommendation does not proceed to the interview, then the person recommending receives an explanatory e-mail or telephone call from RABK explaining this decision. At the end of this round, each jury member provides either a positive or negative assessment per applicant. In general, applicants with only positive scores from all jury members are invited to the interview.

The interview round consists of applicants assessed by two interdisciplinary jury groups – each group consisting of three renowned visual artists and a facilitator – in a 30-minute interview per group with the applicant. The two facilitators take detailed notes during the interviews. Although RABK does not use formal selection criteria, we were able to determine 11 distinct criteria, based upon extensive analysis of jury notes (discussed in detail in the next section). We found that in the interview round, focus is placed upon 11 distinct criteria and the final jury assessment, covering items such as: developmental potential, i.e., technical and creative processes and connection between art practice and theory as well as individual characteristics, i.e., communication skills and ability to accept critique. After the interview, each group give their assessment, i.e. positive or negative, based on (sometimes) extensive discussions about the applicant[15]. The final admission decision is made collectively after both group evaluations have been made. This final decision does not follow mechanically from the jury group assessment, as other cohort dynamics are also considered, i.e., nationality, artistic disciplines and gender.

2.4 DATA

2.4.1 Data Collection

We have two sets of data. The first set contains data collected from the application procedure up to the interview round, these data include: age, gender, country of origin, formal art education or not, whether applied (and rejected) before, and letter of recommendation. We will call this biographical data.

The second data set, called jury data, derives from the notes taken during the interview round for applicants to the RABK in 1995, 1997 and 1999 up to and including 2008. We focus on the interview round, where annually 60-65 applicants were invited, for a few reasons: First, the interview round has been the final round throughout our observation period inviting approximately the same amount of applicants every year. Second, this round is exactly documented. Third, selection, compared to the other rounds, is most rigorous in this round. Fourth, multiple selection criteria are implicitly used to judge applicants' potential quality. And fifth, we test if receiving excellent scores on the multiple selection criteria compared to other applicants results in admission. This enables us to assess the predictive value of this selection process using multiple selection criteria on not only admission but also on future career performance for both accepted and rejected applicants. In addition to the interview notes, our data contains information from three different sources: interviews with jury members and non-participatory observation to distill the exact selection criteria and weight of each individual criterion.

2.4.2 Interviews and non-participatory observation

As the RABK does not have an official list of selection criteria, non-participatory observation and interviews with jury members provided a deeper understanding of the selection procedures, especially in the selection and interview rounds, by helping to identify exact criteria used to assess the applicants and the weight of importance per criterion in each category in these rounds.

Over a four-month period, one of the authors observed two-hours of each type of round[16]; i.e., (pre)-pre selection, pre-selection, selection, and interviews. After each selection and interview round, semi-structured individual meetings with the jury members took place. Each meeting lasted approximately 15 minutes and was designed so that each jury member could rate the importance of various selection criteria that were observed during that round. After these interviews took place, an in-depth interview was conducted with the head of the residency program, who was also a facilitator in the interview round during the past ten years, to assess the accuracy of the multiple selection criteria and to evaluate the jury members' ratings of these criteria. In general, there was agreement among the jury members' rankings and the assessment of these rankings by the head of the residency program; thus no changes were made. Table 1 (see Appendix A) shows the multiple selection criteria with the assigned weight in percentage and importance. A score of one indicated lesser importance, while a score of five reflects great importance in either the selection or interview round.

There is a high degree of continuity in the composition of the jury, with only 9 out of 30 members serving once and 21 members serving in more than one year, and the list may be taken to be valid for our entire interval of observation.[17]

2.4.3 Analysis of interview notes

All interview notes have been transcribed *verbatim*, and three raters scored these notes based upon multiple selection criteria observed during the non-participatory observation and interviews with jury members. Of the total number of applicants accepted in our observation period (N=299) we have interview notes for 53% (N=157), of the total number of applicants rejected in our observation period (N=398) we have interview notes for 50% (N=198).

14 The shortlists are not kept on file.
15 We did not distinguish cases with jury notes from one or from two jury groups: all scores were averaged.
16 This was the first time RABK had allowed data collection during their selection process. To keep possible interference with the process at a minimum, RABK specified time allocations for non-participatory observation and meetings with selection committee members.
17 Among the 30 multiple jurors, the full distribution during our observation period: 3 jurors served on 11 selection committees, 2 jurors on 8 selection committees, 5 jurors on 6 selection committees, 2 jurors on 5 selection committees, 2 jurors on 4 selection committees, 4 jurors on 4 selection committees, 3 jurors on 2 selection committees and 9 jurors on 1 selection committee. Moreover, the same facilitator was present.

We used an ICC model[18] to measure the consistency of agreement between the three raters scoring the RABK interview notes on 11 selection criteria and the final jury assessment. Specifically, a two-way random effects ICC model was used and the consistency of absolute agreement between raters per criterion was measured. Absolute agreement refers to the correlation between measurements of the criteria and it is also a ratio of the between target variance to the total variance of measurements in a two way random effects model (Shrout and Fleiss, 1979). The total variance is the sum of the between target, between rater and error variance. The unit of analysis is the mean rating across the three raters, which is useful to use when a group of raters score the same criteria (Shrout and Fleiss, 1979). Our measure of absolute agreement-ICC is denoted as ICC (2,3), meaning we have 3 raters who each rate the same data. The literature concedes that an ICC value of 0.70 or higher is acceptable.

The three raters who scored these data were two master students in Business Administration and one of the authors. These two student raters were chosen because of their general interest in contemporary visual art; namely one rater's mother is a visual artist, and the other rater has taken university level courses in art history, focusing on modern and contemporary art. The two students received a detailed three-hour training session from the third rater about the 11 selection criteria and final jury assessment, and how these criteria were extracted from the non-participatory observation, interviews with jury members and final in-depth interview with the head of the residency program. Additionally, the training session provided the students background information on the various selection rounds at the RABK, particularly the last round – the interview. The selection criteria were explained in detail and a matrix was provided specifying each criterion with examples. Additionally, art specific context and terms were explained. After scoring 100 interview notes based on the selection criteria, the raters met to discuss questions that arose during scoring of the jury notes. These meetings were repeated twice. The students did not receive compensation, but used the collected data as an empirical basis to write their master thesis.

The 11 selection criteria and final jury assessment used to score the interview notes are based on the individual's developmental potential, connection between art practice and theory and individual characteristics, which have been sub-categorized as: content, context, craft, creative, expertise, critique, communication, career fit, collaboration, contradiction, personality, and final jury assessment. The raters read the interview notes and then assigned a value of 1, 2 or 3 to the selection criterion they thought the interview notes addressed. For most criteria, a value of 1 represented a negative jury assessment, a value of 2 represents neutral assessment, and a value of 3 represented a positive jury assessment. With assessing the value 2 if there is no explicit comment in the notes on this criterion, we implement our interpretation that if nothing is said, the candidate apparently scores neither clearly positive nor clearly negative on this criterion.

18 Intra class correlation coefficient (ICC) is used to measure the consistency or homogeneity of measurements given to the same target by different raters (Shrout and Fleiss, 1979).

For two criteria the scales were slightly different, namely: for *contradiction,* a value of 1 represented complete or partial disagreement between jury members, and a value of 3 represented total agreement. In three records, there was no score on this criterion; we assigned a value 3 as silence can express agreement but not disagreement[19]. For *final assessment* a value of 1 represented an overall negative evaluation and a value of 3 represented an overall positive evaluation, while 2 represented a missing value[20].

TABLE 2.1 Description of selection criteria

- Work related ***content*** refers to the development of the artist/artworks over time.
- ***Context*** refers to the artists' understanding of art historical references.
- ***Craft*** refers to the technical aspects of the artwork/ ability in actually making the artwork.
- ***Creativity*** refers to the innovative aspects/ novel contribution of the artwork.
- ***Expertise*** refers to art education or combinations with other academic disciplines, e.g., biology, architecture.
- ***Critique*** refers to how open the artist is to criticism and change.
- ***Communication*** refers to the artists' ability to communicate about the artworks, i.e., reason for artwork.
- ***Fit*** refers to the stage of the artists' career and the fit within the RABK community.
- ***Collaboration*** refers to if jury member(s) would like to have the artist as a colleague.
- ***Contradiction*** refers to the level of opposing views between jury members about an applicant.
- ***Personality*** refers to if the jury liked the personality of the applicant.
- ***Final assessment*** refers to jury members' overall assessment of the applicant.

On average, the raters were consistent with one another in their measurements and the *average scores* across all criteria resulted in ICC(2,3) 0.866. Raters also consistently agreed on the jury assessment of the applicants *fit* ICC(2,3) 0.816, *collaboration* ICC(2,3) 0.789 and *final jury assessment* ICC(2,3) 0.787. On the level of the other criteria, the measurements were slightly lower; namely, *content* (.561), *context* (.502), *craft* (.657), *creative* (.645), *expertise* (.582), *critique* (.660) *communication* (.685), *contradiction* (.541) and *personality* (.678). These somewhat inconsistent measurements could be in part due the ambiguous comments made during by the jury. For instance, one applicant received the following comment: "beautifully new paintings compared to last year". One rater interpreted this comment as being a positive comment about the artist's *craft* (technique), the other about *creativity* (innovation), and the third about *content* (development) and *creativity*.

19 We used regressions where we assigned a value 1 rather than 3; this yielded similar results.
20 In 8 observations, a final jury assessment score was missing; all these applicants were rejected. We conducted a regression analysis excluding these observations; this yielded similar results.

2.4.4 Performance measures

According to the RABK mission statement: *RABK aims to develop talent in visual arts. The Academy selects and facilitates exceptional talent and offers selected artists a platform for further development and deepening of the work of art. Exhibition of the work and access to the international network are included in the two-year residency program. For many alumni the internship at RABK has led to their international breakthrough (www.rijksakademie.nl).* We interpret this goal as selection of the best artists from the large pool of applicants, and the facilitative role of RABK in their development to achieve leading positions, i.e., success, in the contemporary art world on an international level. We measure success by two variables: auction sales (from Artnet.com) and reputational rankings based upon exhibition history (from ArtFacts. Net.).

Our first variable of interest is career level sales made at public art auctions. We retrieve this information from Artnet.com, which was founded in 1989. Artnet.com is an art market website that provides detailed information about more than 9 million public art auction results from 1,600 international auction house (Artnet.com annual report, 2014). Our second variable of interest is reputational rankings. We access this information from ArtFacts.Net, which was established in 2001. ArtFacts.Net is a website that ranks contemporary visual artists based upon their annual exhibitions at galleries and museums worldwide. Their aim is "to provide an online structure for art institutions and Internet users, and offer insight into the current state of the contemporary art world" (ArtFacts.Net, 2005). ArtFacts.Net provides reputational rankings for over 100,000 visual artists, based upon annual exhibitions at more than 600,000 galleries, museums and other venues worldwide. Artists are allocated points[21] based on whether or not the exhibition was solo or group at private galleries or public institutions or biennales and other regular exhibitions, and the geographic location, with art centers such as London and New York City receiving more points than Des Moines, Iowa. The artists with the highest number of points are given the lowest rank, e.g., the top ten ranked artists in 2016, in ascending order are: Andy Warhol, Pablo Picasso, Bruce Nauman, Gerhard Richter, Joseph Beuys, Cindy Sherman, John Baldessari, Lawrence Weiner, Sol LeWitt and Ed Ruscha. ArtFacts.Net data has been widely used as a measure of success in management science literature (Ertug, Yogev, Lee, Hedstrom, 2016) and the literature of cultural sociology (Velthuis, 2013; Yogev and Grund, 2012).

2.5 AT FIRST GLANCE: THE ADMITTED AND THE REJECTED

2.5.1 Summary Statistics

In Table 2.2, we present basic statistics of our data. Some 43 % of the applicants in our sample have been accepted, almost half are female; sculpture, video and painting take up roughly equal shares; drawing and photography are less frequent. More than half of the applicants invited to the interview round were European: 28% Dutch nationals or have lived in the Netherlands longer than one year and 36% from other Western and Central European countries.

21 For a complete explanation of ArtFacts.Net point allocation, see http://faculty.georgetown.edu/irvinem/visualarts/ArtMarket/ArtFacts.Net-ArtistRanking-Forumula-2005.pdf.

The other 36% of the applicants originate from: North America 7%, South and Central America 8%, India, Asia, and Australia 10% and Africa, and Middle East 11%.

TABLE 2.2 Sample characteristics, all applicants with and without jury notes

	All	*Jury notes*	*No jury notes*
Sample size	697	355	342
Percentage accepted	42.90	44.22	41.52
Percentage female	48.20	52.68	43.57
Percentage letter of recommendation	18.94	20.29	17.54
Percentage no previous art education	4.59	3.66	5.55
Discipline:			
Sculpture	27.69	33.52	21.64
Video/film	23.96	27.32	20.47
Painting	28.55	22.82	34.50
Drawing/graphic	8.03	6.76	9.36
Photography	11.76	9.58	14.03
Percentage with 2013 ArtFacts.Net rank	69.01	69.86	68.13
Percentage with 2015 ArtFacts.Net rank	71.30	71.55	71.05
ArtFacts.Net ranking, mean 2013-2015	31776.79	32611.70	30878.37
Percentage with either 2013 or 2015	10.10	11.98	8.06
Percentage with no ArtFacts.Net rank in observation period	26.11	24.78	27.48
Percentage any sales	12.91	12.96	12.86
Mean number of sales if positive	5.59	5.04	6.16
Mean sales price in euros of sales if positive	167552.10	214052.30	118938.30

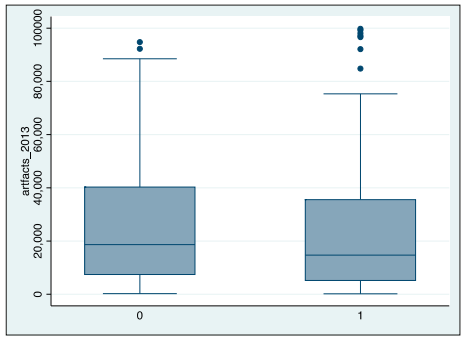
The results in Table 2.2 suggest that applicants without jury notes are the marginally weaker group. They are accepted at a slightly lower frequency, a smaller share has a letter of recommendation, a larger share does not have formal art school training, and more individuals from this group have not been ranked by ArtFacts.Net compared to the group with jury notes. However, of those that were ranked by ArtFacts.Net in 2013 and 2015, the group without notes has slightly better rankings. Additionally, these applicants - if they had any sales at all- had more sales, although the mean sales price is considerably lower compared to the group with jury notes. We will return to the issue of selective availability of jury notes below.

2.5.2 Preliminary Analyses

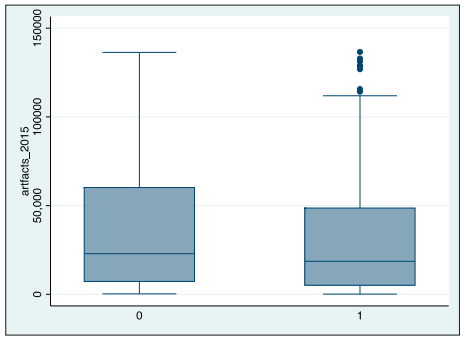
As a first glance about the effectiveness of selective entry, we can consider the performance measures for the accepted applicants and the rejected ones. Figure 2.1 gives the distribution of 2013 and 2015 ArtFacts.Net scores and the distribution of the individual artists' median scores over both years[22], for accepted (d =1) and rejected (d=0) applicants. The top horizontal lines in Figure 2.1 show the maximum, 3rd quartile, median, 1st quartile and minimum score, respectively. A dot represents an individual outlier. Accepted applicants perform better, with lower average rank number (i.e. better ranking) and with a shorter lower tail: the bottom performance quartile scores much better ranking positions. Selection appears to cut off the low end of the distribution. A comparison is made between ArtFacts.Net ranking 2013-2015 average scores with and without the outliers, although no real difference in distribution is observed when the outliers are removed. In all boxplot distributions, individuals without Artfacts.Net scores are excluded.

FIGURE 2.1 Boxplot distributions of ArtFacts.Net rankings

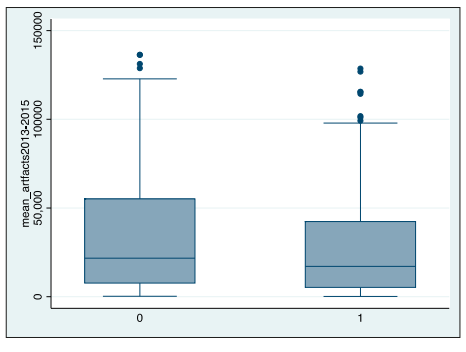
ArtFacts.Net rankings 2013



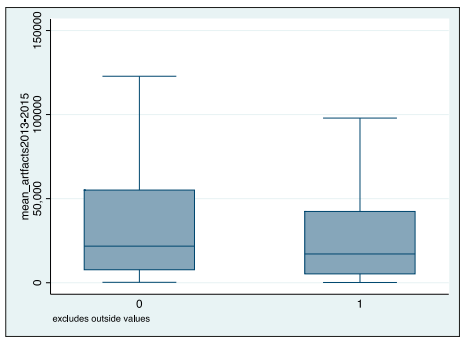
ArtFacts.Net rankings 2015



ArtFacts.Net ranking 2013-2015 average scores (including outliers)



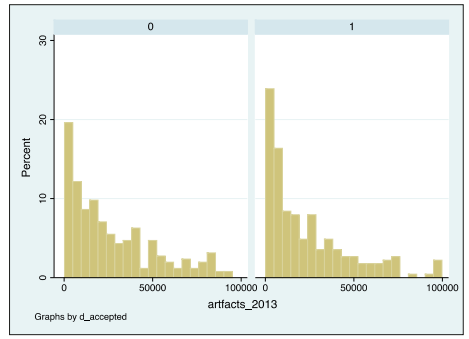
ArtFacts.Net ranking 2013-2015 average scores (excluding outliers)



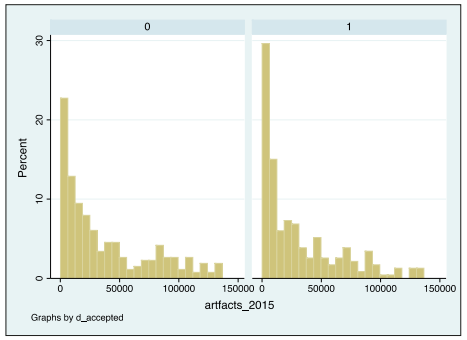
22 For those individuals who were ranked in either 2013 or 2015, their average rank is simply the rank they received in that particular year.

FIGURE 2.2 Histograms of ArtFacts.Net rankings

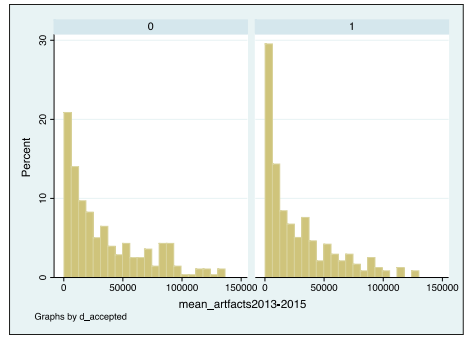
ArtFacts.Net rankings 2013



ArtFacts.Net rankings 2015

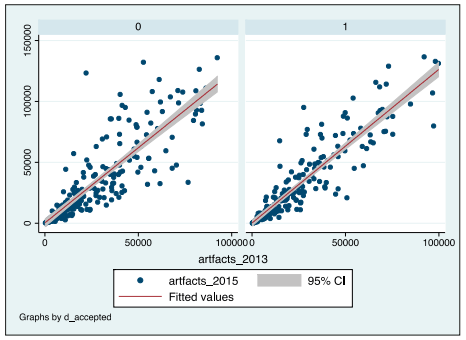


ArtFacts.Net ranking 2013-2015 average scores



The histograms in Figure 2.2 confirm that the distributions differ. A formal t-test on the means reveals that 2013 ranking (mean of 23,496.5 for the accepted and 26,798.33 for the rejected, at standard deviations of 23,711.26 and 24,224.91, respectively) is not significantly different from zero at $p<.05$.

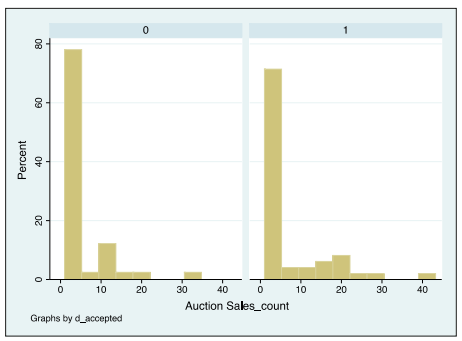
FIGURE 2.3 Two-way scatter plots for ArtFacts.Net rankings



The 2015 ranking (mean of 31,064.9 for the accepted and 37,420.13 for the rejected, at standard deviations of 33,585.12 and 37,468.61, respectively) is significantly diffe-rent from zero at p<.05; and the average 2013-2015 ranking (mean 28,242.76 for the accepted and 34,789.61 for the rejected, at standard deviations of 29,492.08 and 33,244.62, respectively) is also significantly different from zero at p<.05.

The two-way scatter plots in Figure 2.3, including the confidence intervals around the predicted values, show a positive linear relationship between rankings in both years for accepted and rejected applicants.

FIGURE 2.4.1 Distribution of number of artworks sold if positive for accepted (1) and rejected (0) artists.



We have also compared sales performance in Figure 2.4.1. Among the accepted, 49 artists (16.39%) who had any auction sales and among the rejected 41 artists (10.30%). The 49 accepted artists with any sales sold 319 works, the 41 rejected artists who sold any work sold 184 artworks, i.e., if artists sold artworks, the average number of works sold per individual was 6.5 for the accepted and 4.5 for the rejected. The average selling price for the accepted candidate who sold any work was 18,616 euros (SD 130,479 euros) and 23,139 euros (SD 389,885 euros) for the rejected who sold any work at all.

FIGURE 2.4.2 Distribution of number of works sold (log) if positive for accepted (1) and rejected (0) artists.

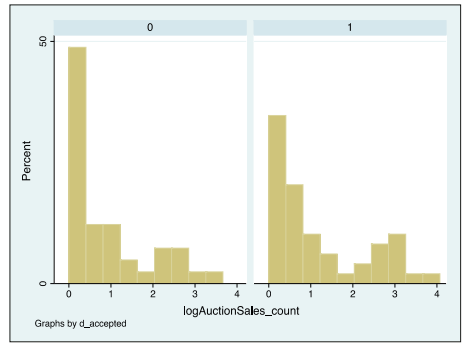


Figure 2.4.2 shows the distributions of number of works sold if positive, for the accepted and the rejected. A formal t-test on the means reveals that the mean sales count (6.510 for the accepted and 4.487 for the rejected, at standard deviations of 8.646 and 6.395, respectively) are not significantly different from zero at p<.05; and the mean sales amount in euros (113,598.30 for the accepted and 232,033.50 for the rejected, at standard deviations of 43957.53 and 188,779.20, respectively) are not significantly different from zero at p<.05. Additionally, we conducted a chi-square test of independence to test for equality of proportions between the accepted and rejected applicants and the number of sales at auctions; we found that the proportions are not significantly different from each other at p<.05.

FIGURE 2.4.3 Distribution of sales price (log) of works sold if positive for accepted (1) and rejected (0) artists.

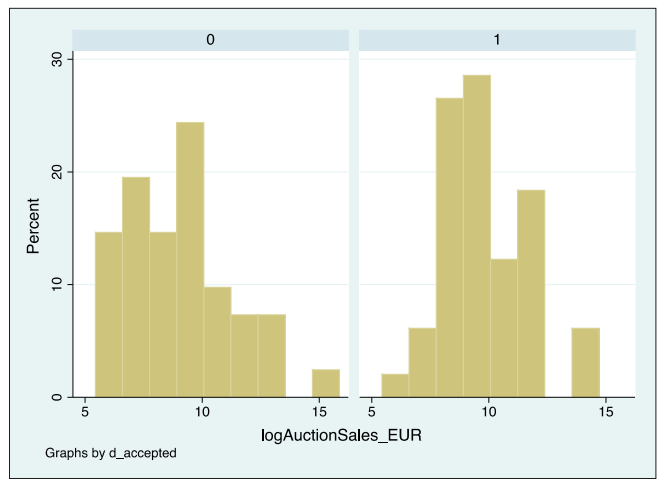


TABLE 2.3[23]: Residuals and probability of acceptance (probit)

	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*
res_l~2013	-0.083**						0.069
	(0.034)						(0.166)
dartf~2013	-0.913*						1.217
	(0.527)						(1.668)
res_loga~5		-0.085**					-0.157
		(0.028)					(0.146)
dartf~2015		-0.904**					-1.712
		(0.435)					(1.376)
r~20132015			-0.027**				0.031
			(0.007)				(0.234)
d_2015_13			-0.496**				-0.085
			(0.192)				(1.707)
d_no_rank				-0.311**			0.000
				(0.111)			(.)
res_logA~R					0.035**		0.001
					(0.015)		(0.022)
res_Auct~t						0.032**	0.017
						(0.015)	(0.021)
N	697	697	697	697	697	697	697
pseudo R2	0.018	0.022	0.019	0.008	0.006	0.005	0.027

Standard errors in parentheses
* p<0.10, ** p<0.05

In Table 2.3 we consider the probability of acceptance in relation to performance measures, ignoring the reverse effect that a residency at the RABK may have on later performance (see Appendix C for description of variable names). If we assume that performance is exogenous in this respect, that selection aims to admit the "better" artists and that the quality the jury aims for is reflected in our performance measures, regression of admittance on performance would reveal the success of the jury decisions.

As older cohorts of applicants have had more time to realize their potential, we control for "career age" by measuring performance as the residual from a regression on career age and its square. The results in Table 2.3 indicate that realized future success, i.e., a lower ArtFacts.Net ranking or higher sales made at auction (count and euro amount) has a significant effect on probability of admission.

23 We created two dummy variables indicating if an individual had an ArtFacts.Net ranking in 2013 or in 2015. If that was true then a value of one was assigned, otherwise a value of zero was given. We also made a dummy variable indicating if an artist was not ranked in 2013 or 2105 then a value of one was given, otherwise zero.

When all variables are included in model 6, the separate effects are diluted to insignificance. For those artists who received an ArtFacts.Net rank in 2013 and/or 2015, the correlation between the residuals of the variables ArtFacts.Net rank 2013 and ArtFacts.Net rank 2015 is r=0.84 and the correlation between the residuals of sales count and sales price is r=0.69. We may conclude that artists who turn out to perform better had a greater chance of having been being admitted to the RABK.

This preliminary analysis indicates differences between accepted and rejected applicants, but the differences are not always statistically significant. The probability of admittance appears to have been higher for artists who later secure higher reputation ranking and who sell more artworks, either as number of sales or price paid. Conversely, accepted artists have higher mean ranking than rejected artists, but this difference is not statistically significant in 2013. More often they have any positive sales but they realize a lower selling price. They sell more works, but the difference is not statistically significant. Visually, selection appears to cut off the bottom tail of the ranking distribution; by consequence, the ranking distribution of the accepted has smaller standard deviation than the distribution of the rejected.

2.6 SELECTION AND THE IMPACT ON PERFORMANCE

2.6.1 Interview notes and selectivity

To analyze the admission decision, we will use the two types of explanatory variables, the biographic data and the jury data, consisting of the scores on 11 multiple selection criteria and the final jury assessment. We make this distinction, as we are particularly interested in the effectiveness of the jury assessments. As we do not have jury notes for each applicant, we will first consider if availability of notes is selective rather than random. In a probit regression of having jury notes we do not find an effect on the acceptance decision nor any of the performance measures. Table 2.4 we show results when we add the biographic controls. This tells the same story: having jury notes or not having notes is not related to the acceptance decision or to performance measures. There are, however, differences among disciplines (higher probability for video and sculpture) and by gender: a higher probability of notes for women. We have no explanation for these findings[24]. These results confirm that the availability of jury notes is random, and therefore, have no effect on selectivity in relation to individual quality or qualification.

24 Gender composition of the jury: Of the 30 individuals who served as jurors on the selection committee(s), 40% were women (n=12), and 60% were men (n=18).

TABLE 2.4 Probability of interview notes related to performance (probit)

	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*
d_accepted	-0.077							-0.071
	(0.127)							(0.130)
d_ref_le~r	-0.119	-0.119	-0.120	-0.125	-0.123	-0.128	-0.125	-0.110
	(0.163)	(0.163)	(0.163)	(0.163)	(0.163)	(0.163)	(0.163)	(0.164)
d_autodi~s	-0.047	-0.065	-0.067	-0.052	-0.053	-0.055	-0.054	-0.105
	(0.302)	(0.303)	(0.303)	(0.305)	(0.302)	(0.302)	(0.302)	(0.314)
d_sculpt~n	0.905**	0.881**	0.893**	0.892**	0.900**	0.892**	0.894**	0.892**
	(0.168)	(0.168)	(0.168)	(0.168)	(0.168)	(0.169)	(0.168)	(0.172)
d_videof~m	0.869**	0.833**	0.843**	0.853**	0.859**	0.849**	0.856**	0.839**
	(0.183)	(0.184)	(0.184)	(0.182)	(0.182)	(0.184)	(0.182)	(0.190)
d_female	0.288**	0.301**	0.294**	0.294**	0.293**	0.289**	0.293**	0.302**
	(0.125)	(0.126)	(0.126)	(0.126)	(0.126)	(0.125)	(0.125)	(0.127)
d_nation~y	-0.155	-0.134	-0.158	-0.151	-0.160	-0.166	-0.155	-0.124
	(0.145)	(0.147)	(0.146)	(0.146)	(0.144)	(0.145)	(0.146)	(0.148)
age_inte~w	-0.098	-0.113	-0.101	-0.106	-0.101	-0.099	-0.104	-0.110
	(0.115)	(0.122)	(0.118)	(0.120)	(0.117)	(0.117)	(0.119)	(0.123)
age_2	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
d2x_inte~w	0.144	0.139	0.150	0.144	0.153	0.149	0.147	0.146
	(0.204)	(0.203)	(0.203)	(0.204)	(0.203)	(0.203)	(0.203)	(0.207)
d_drawin~s	0.310	0.291	0.294	0.303	0.308	0.296	0.305	0.282
	(0.240)	(0.240)	(0.240)	(0.240)	(0.241)	(0.241)	(0.239)	(0.246)
d_photog~y	0.013	-0.018	-0.001	0.002	0.009	0.000	0.006	-0.056
	(0.219)	(0.221)	(0.220)	(0.220)	(0.219)	(0.220)	(0.219)	(0.225)
res_l~2013		-0.047						-0.193
		(0.060)						(0.217)
dartf~2013		-0.619						-2.755
		(0.937)						(2.308)
res_loga~5			-0.026					0.174
			(0.051)					(0.265)
dartf~2015			-0.411					1.809
			(0.791)					(1.989)
r~20132015				-0.004				-0.045
				(0.009)				(0.385)
d_2015_13				-0.021				-0.204
				(0.241)				(2.821)
res_logA~R					0.004			0.009
					(0.019)			(0.028)
res_Auct~t						-0.005		-0.015
						(0.019)		(0.027)
d_no_rank							-0.046	0.000
							(0.140)	(.)
N	697	697	697	697	697	697	697	697
Cohort yr effects	yes	yes	yes	yes	yes	yes	yes	yes
pseudo R2	0.429	0.429	0.429	0.428	0.428	0.428	0.428	0.434

Standard errors in parentheses
* p<0.10, ** p<0.05

In the preliminary analysis we also conducted a principle components analysis (see Appendix F) but we were unable to give meaningful interpretations to the groupings. In fact, the correlations among the criteria are rather low, with two sets of variables content and creativity and context and communication having the highest correlations at r= 0.26 and r= 0.22, respectively. As reduction of dimensionality is attractive, we did aggregate variables into three groups, based on similarity or closeness of the criteria, in our own reading of the definitions:

Group name	*Selection criteria*
(1) Art quality	Craft and creativity
(2) RABK goals	Fit, context, critique
(3) Individual quality	Personality, contradiction, collaboration, communication, personality, expertise

The rational to organize the multiple selection criteria into these specific sub-groups is as follows: If the jury process works as the RABK intends, to find promising artists that fit the RABK setting and will benefit most, group (2) should have a substantial weight in the admission decision. It may also be that the jury, as many juries do, simply focuses on perceived quality. In that case, group (1) should dominate the decision. If group (3) dominates, it's just the impression that an artist makes on the jury: nice person, good communication, sells him/herself well.

2.6.2 Multiple selection criteria

In Table 2.5, we analyze the acceptance decision with a probit regression. We start with biographic data only and then add data from the jury assessment in several steps. With only biographic data, on the full sample, pseudo R^2 is low. Dutch applicants and video artists have a higher probability of admission in the full sample. As we can only analyze the effects of the jury assessment if we have jury notes, we repeat this analysis for the subsample with notes. Now, in addition, sculpture artists are also significantly more often admitted, a letter of reference has a significant positive effect and having an interview history is strongly negative. The dummy for the years with a different jury procedure (1995-1998) is only significant when we add the individual selection criteria and the groupings; along with this control variable, we have also included cohort year effects.

In model 3 we add the variables for the individual criteria, and we see that four of them have a significant positive effect. Ranked by size of the coefficient, they are: collaboration, creativity, fit and content. This is a deviation from the weights of the criteria on the final assessment. If, instead of the separate criteria, we use the final assessment, the pseudo R^2 increases. Model 6 shows that significance of a criterion in the linear specification does not always point to a situation where successive scores per criterion have monotonic effect on the admission decision.[25] This is reiterated in a specification (not reported) where we changed the reference level of the jury criteria, now with the lowest level as base, rather than the middle level.

25 We have considered if criteria with larger variance of scores among individuals (indicating more perceived heterogeneity) are better predictors of admission. This is not the case; there is no relation between score variance and t value of criteria (see Appendix D).

Creativity has its significant effect mostly because of a positive effect for scoring above average, while personality and fit has a negative effect when scoring below average. When we add the three groups in model 7, the pseudo R^2 is lower than when we only include the residual for final assessment. In this specification, artistic quality has the strongest positive effect, followed by fit with the RABK goals; personality has no significant effect.

We conclude that among the biographic data, nationality and a reference letter have a significant positive effect on the admission decision and that video artists have a greater likelihood of admission than artists from other disciplines, while repeated participation in the jury process has a strong negative effect. Among the separate jury criteria, significant monotonic effects are not the rule. Scoring above average on creativity has a clear positive effect, scoring below average on personality and fit has a clear negative effect. The final jury assessment, combining much information collected in the jury process, on its own has a strong effect on the admission decision. Averaging the jury assessment into three variables shows that artistic quality and fit to the RABK goals have significant effect on the admission decision, while personality does not.

TABLE 2.5 Probability of acceptance (probit)

	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*
age_inte~w	-0.048	-0.007	0.024	-0.043	-0.024	-0.021	0.040
	(0.064)	(0.083)	(0.104)	(0.088)	(0.122)	(0.120)	(0.097)
age_2	0.000	0.000	-0.000	0.001	0.000	0.000	-0.001
	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)
d_female	-0.119	-0.089	-0.098	-0.215	-0.192	-0.200	-0.102
	(0.100)	(0.141)	(0.152)	(0.156)	(0.172)	(0.177)	(0.148)
d_nation~y	0.197*	0.300*	0.378**	0.593**	0.756**	0.789*	0.393**
	(0.111)	(0.167)	(0.183)	(0.186)	(0.214)	(0.227)	(0.178)
d_autodi~s	0.179	0.093	0.344	0.360	0.671	0.866	0.431
	(0.238)	(0.376)	(0.441)	(0.388)	(0.466)	(0.491)	(0.425)
d_ref_le~r	0.097	0.336*	0.322*	0.358*	0.374*	0.442*	0.347*
	(0.128)	(0.182)	(0.195)	(0.198)	(0.220)	(0.230)	(0.190)
d2x_inte~w	-0.208	-0.637**	-0.664**	-0.986**	-1.142**	-1.23**	-0.630**
	(0.160)	(0.252)	(0.283)	(0.277)	(0.321)	(0.337)	(0.272)
d_sculpt~n	0.202	0.344*	0.378*	0.315	0.369	0.460*	0.314
	(0.134)	(0.205)	(0.221)	(0.223)	(0.248)	(0.260)	(0.213)
d_videof~m	0.447**	0.421**	0.462**	0.389*	0.500**	0.591*	0.435**
	(0.140)	(0.205)	(0.223)	(0.224)	(0.251)	(0.269)	(0.215)
d_drawin~s	0.182	-0.117	-0.319	-0.022	-0.201	-0.221	-0.296
	(0.196)	(0.320)	(0.349)	(0.336)	(0.379)	(0.395)	(0.335)
d_photog~y	0.187	0.293	0.362	0.269	0.355	0.286	0.383
	(0.172)	(0.276)	(0.297)	(0.305)	(0.336)	(0.347)	(0.285)
d_jur~1998	0.262	0.361	0.686*	0.005	0.378	0.271	0.655*
	(0.222)	(0.313)	(0.352)	(0.351)	(0.421)	(0.431)	(0.341)
CONTENT			0.207**		0.212**		
			(0.096)		(0.106)		
low score						-0.461	
						(0.285)	
high score						-0.026	
						(0.321)	
CONTEXT			0.018		-0.022		
			(0.137)		(0.152)		
low score						0.042	
						(0.211)	
high score						-0.094	
						(0.398)	
CRAFT			0.145		0.171		
			(0.114)		(0.128)		
low score						-0.067	
						(0.231)	
high score						0.382	
						(0.280)	
CREATIVITY			0.444**		0.553**		
			(0.100)		(0.113)		
low score						-0.180	
						(0.311)	
high score						1.033**	
						(0.335)	
EXPERTISE			-0.120		-0.222		
			(0.171)		(0.190)		
low score						0.034	
						(0.233)	
high score						-0.984*	
						(0.523)	
CRITIQUE			0.192		0.239		
			(0.154)		(0.171)		
low score						-0.227	
						(0.232)	
high score						0.426	
						(0.396)	
COMMUNICATION			-0.050		-0.054		
			(0.125)		(0.139)		
low score						0.128	
						(0.221)	
high score						0.041	
						(0.340)	
FIT			0.415**		0.451**		
			(0.117)		(0.133)		
low score						-0.577*	
						(0.223)	
high score						0.305	
						(0.282)	
COLLABORATION			0.545**		0.622**		
			(0.244)		(0.273)		
low score						-0.717*	
						(0.391)	
high score						0.678	
						(0.483)	
CONTRADICTION			-0.034		-0.050		
			(0.079)		(0.089)		
low score						0.190	
						(0.189)	
high score						omitted	
PERSONALITY			0.192		0.207		
			(0.137)		(0.145)		
low score						-0.663**	
						(0.250)	
high score						-0.221	
						(0.263)	
resFINAL				0.807**	0.868**	0.877**	
				(0.094)	(0.100)	(0.095)	
Art quality							0.768**
							(0.149)
RABK Goals							0.681**
							(0.214)
Individual quality							0.075
							(0.254)
N	697	355	355	355	355	355	355
Cohort yr effect	yes	yes	yes	yes	yes	yes	yes
pseudo R2	0.034	0.063	0.187	0.319	0.371	0.393	0.153

Standard errors in parentheses
* p<0.10, ** p<0.05

2.6.3 Switching regressions

We estimate switching regressions, also known as the Roy model, to measure the performance of the accepted and rejected applicants. Switching regressions are similar to two-step selection models (Heckman 1974,1976), i.e., to get consistent estimates, the mean zero restriction on the error term, conditional on the explanatory variables, is restored by including an estimate of the selection bias. This means that these models are efficient as long as the functional relationship between the outcome equation and the selection equation are correctly specified.

We first estimate the parametric two-step selection models, which include an outcome equation and a selection equation for two distinct performance outcomes: sales at public auctions and reputational ranking. For sales, we have complete records: any artist who ever sells anything in an auction is recorded. We transformed amount of sales in euros to the natural log, adding 1 euro if sales were zero. The range of sales in euros expressed in natural log is: 15.86 (for the best performer) and 5.44 (for the worst); we observe n=355 for both performance and selection equations for accepted and rejected. For reputational rankings, measured by ArtFacts.Net, we took the mean of 2013 and 2015 rankings if both positive, or one year ranking if only one positive; otherwise we assigned those individuals who were not scored in either 2013 or 2015 a score of 25 (log), which is an artificially high score, roughly 10^11. The range of ArtFacts.Net scores expressed in natural log is: 5.01 (for the best performer) and 11.82 (for the worst); again, we observe n=355 for both performance and selection equations for accepted and rejected. We also include a dummy variable coded 1 if the artist has never been ranked, otherwise 0.

In the outcome equation, we estimate the relationship between both performance outcomes and the explanatory variables, i.e., biographic data and multiple selection criteria. We also specify the variable career age, which is measured as 2015 minus year applied. We include this variable to control for tenure, i.e., artists with longer careers had opportunities to sell more artworks and participate in more exhibitions compared to artists with shorter career trajectories. In the selection equation, we estimate the binary decision of acceptance and rejection, using the same explanatory variables as in this first equation. However, in this equation we also include two instruments as variables to enhance identification of these models. The first instrument is a set of dummy variables indicating artists' origin based on geographic region.[26] We include these variables because RABK explicitly selects some applicants from under represented geographical regions.[27] The second instrument is a dummy variable indicating if the applicant is Dutch, coded as 1, otherwise 0. We make this distinction because RABK categorizes artists living in the Netherlands for at least one year, regardless of their citizenship, as being Dutch. In the selection equation we do not include the variable career age or the variable measuring no Artfacts.Net rank, as these two variables do not have an effect on selection.

Similar to parametric two-step selection models, semi-parametric selection models consist of performance and selection equations, where the functional relationships between the two are specified.

26 Nation_1 represents Europe, Nation_2 represent North America, Nation_3 represents South America, Mexico and the Caribbean, Nation_4 represents Asia, and Nation_5 represents Africa and Middle East (base).

27 This information was provided during an interview with the RABK director in 2009.

We estimated Newey (1999) semi-parametric selection models with polynomial approximations of orders 2, 3 and 4. The estimation results from both the parametric and semi-parametric two-step selection models are quite similar; with the polynomial approximation of order 4, in the semi-parametric selection models, indicating selectivity with respect to the accepted candidates but not to the rejected ones, similar to our results in the parametric estimations. This result leads us to conclude that the assumption of normality made in the parametric selection models is not too restrictive. Because of the similarity in our results, we will continue reporting results from the parametric two-step selection models. (Results from the semi-parametric selection models available upon request)

TABLE 2.6.1 Explaining reputational rankings by selection criteria

	(accepted)	*(rejected)*
Outcome equation		
career_age	0.035	0.017
	(0.040)	(0.033)
d_1age_i~w	-0.283	-0.263
	(0.273)	(0.257)
d_2age_i~w	-0.466*	-0.071
	(0.257)	(0.230)
d_female	0.049	0.317*
	(0.203)	(0.189)
d_autodi~s	0.197	-1.328**
	(0.564)	(0.527)
d_ref_le~r	-0.039	0.044
	(0.248)	(0.247)
d2x_inte~w	0.408	0.786**
	(0.475)	(0.347)
d_2015_13	0.137	1.206**
	(0.389)	(0.274)
d_no_rank	15.058**	15.227**
	(0.250)	(0.191)
d_no_sale	1.695**	0.802**
	(0.277)	(0.266)
d_sculpt~n	-0.105	0.050
	(0.367)	(0.353)
d_videof~m	-0.417	0.168
	(0.363)	(0.357)
d_painting	0.915**	0.815**
	(0.405)	(0.362)
d_drawin~s	0.123	0.044
	(0.572)	(0.450)
d_jur~1998	0.400	0.332
	(0.367)	(0.347)
ave_c~nt_i	-0.141	-0.293**
	(0.134)	(0.130)
ave_c~xt_i	-0.045	0.133
	(0.169)	(0.171)
ave_craf~i	-0.113	0.013
	(0.146)	(0.144)
ave_crea~i	-0.237	-0.191
	(0.173)	(0.179)
ave_expe~i	0.052	0.020
	(0.211)	(0.200)
ave_crit~i	0.228	-0.002
	(0.201)	(0.190)
ave_comm~i	0.147	-0.111
	(0.156)	(0.152)
ave_fit_i	-0.329*	-0.264
	(0.180)	(0.176)
ave_coll~i	-0.184	0.108
	(0.315)	(0.318)
av~ction_i	-0.104	0.075
	(0.105)	(0.096)
ave_pers~i	-0.077	-0.268
	(0.186)	(0.167)
res_Final	-0.541**	-0.695**
	(0.237)	(0.225)
Selection equation		
d_nation~y	0.715**	-0.715**
	(0.234)	(0.234)
Nation_1	-0.142	0.142
	(0.298)	(0.298)
Nation_2	0.085	-0.085
	(0.376)	(0.376)
Nation_3	-0.664*	0.664*
	(0.379)	(0.379)
Nation_4	-0.265	0.265
	(0.357)	(0.357)
d_1age_i~w	-0.080	0.080
	(0.239)	(0.239)
d_2age_i~w	-0.149	0.149
	(0.212)	(0.212)
d_female	-0.208	0.208
	(0.168)	(0.168)
d_autodi~s	0.724	-0.724
	(0.459)	(0.459)
d_ref_le~r	0.266	-0.266
	(0.211)	(0.211)
d2x_inte~w	-1.261**	1.261**
	(0.331)	(0.331)
d_sculpt~n	0.057	-0.057
	(0.324)	(0.324)
d_videof~m	0.172	-0.172
	(0.330)	(0.330)
d_painting	0.305	0.305
	(0.336)	(0.336)
d_drawin~s	-0.459	0.459
	(0.417)	(0.417)
d_jur~1998	-0.210	0.210
	(0.259)	(0.259)
ave_c~nt_i	0.227**	-0.227**
	(0.104)	(0.104)
ave_c~xt_i	-0.032	0.032
	(0.152)	(0.152)
ave_craf~i	0.141	-0.141
	(0.126)	(0.126)
ave_crea~i	0.535**	-0.535**
	(0.105)	(0.105)
ave_expe~i	-0.137	0.137
	(0.182)	(0.182)
ave_crit~i	0.361**	-0.361**
	(0.171)	(0.171)
ave_comm~i	-0.048	0.048
	(0.136)	(0.136)
ave_fit_i	0.431**	-0.431**
	(0.131)	(0.131)
ave_coll~i	0.640**	-0.640**
	(0.276)	(0.276)
av~ction_i	-0.055	0.055
	(0.086)	(0.086)
ave_pers~i	0.226	-0.226
	(0.145)	(0.145)
Res_Final	0.840**	-0.840**
	(0.096)	(0.096)
N	355	355
Mills ratio		
Lambda	-0.923*	1.331**
Rho	-0.742	0.982
Sigma	1.243	1.354

Standard errors in parentheses
* p<0.10, ** p<0.05

TABLE 2.6.2 Explaining sales at art auctions by selection criteria

	(accepted)	(rejected)
Outcome equation		
career_age	-0.025	0.040
	(0.020)	(0.025)
d_1age_i~w	-0.145	-0.030
	(0.137)	(0.177)
d_2age_i~w	-0.126	0.034
	(0.129)	(0.157)
d_female	-0.250**	-0.029
	(0.101)	(0.130)
d_autodi~s	-0.429	-0.103
	(0.282)	(0.365)
d_ref_le~r	-0.140	0.107
	(0.124)	(0.172)
d2x_inte~w	-0.174	-0.111
	(0.239)	(0.235)
d_2015_13	-0.003	-0.077
	(0.197)	(0.211)
d_no_rank	0.034	-0.082
	(0.128)	(0.146)
d_no_sale	-9.751**	-8.901**
	(0.141)	(0.203)
d_sculpt~n	0.034	0.065
	(0.184)	(0.244)
d_videof~m	-0.003	-0.105
	(0.181)	(0.247)
d_painting	0.039	0.268
	(0.203)	(0.247)
d_drawin~s	0.082	-0.070
	(0.288)	(0.305)
d_jur~1998	0.298	-0.158
	(0.185)	(0.247)
ave_c~nt_i	0.012	0.072
	(0.067)	(0.090)
ave_c~xt_i	0.075	-0.104
	(0.084)	(0.118)
ave_craf~i	0.010	-0.003
	(0.073)	(0.099)
ave_crea~i	0.188**	0.209
	(0.087)	(0.128)
ave_expe~i	-0.022	-0.148
	(0.105)	(0.138)
ave_crit~i	-0.052	-0.043
	(0.100)	(0.131)
ave_comm~i	0.065	0.066
	(0.078)	(0.105)
ave_fit_i	0.078	0.226*
	(0.090)	(0.123)
ave_coll~i	0.122	0.363*
	(0.157)	(0.219)
av~ction_i	-0.029	-0.010
	(0.052)	(0.066)
ave_pers~i	0.133	0.066
	(0.093)	(0.115)
res_Final	0.162	0.183
	(0.119)	(0.159)
Selection equation		
d_nation~y	0.715**	-0.715**
	(0.234)	(0.234)
Nation_1	-0.142	0.142
	(0.298)	(0.298)
Nation_2	0.085	-0.085
	(0.376)	(0.376)
Nation_3	-0.664*	0.664*
	(0.379)	(0.379)
Nation_4	-0.265	0.265
	(0.357)	(0.357)
d_1age_i~w	-0.080	0.080
	(0.239)	(0.239)
d_2age_i~w	-0.149	0.149
	(0.212)	(0.212)
d_female	-0.208	0.208
	(0.168)	(0.168)
d_autodi~s	0.724	-0.724
	(0.459)	(0.459)
d_ref_le~r	0.266	-0.266
	(0.211)	(0.211)
d2x_inte~w	-1.261**	1.261**
	(0.331)	(0.331)
d_sculpt~n	0.057	-0.057
	(0.324)	(0.324)
d_videof~m	0.172	-0.172
	(0.330)	(0.330)
d_painting	-0.305	0.305
	(0.336)	(0.336)
d_drawin~s	-0.459	0.459
	(0.417)	(0.417)
d_jur~1998	-0.210	0.210
	(0.259)	(0.259)
ave_c~nt_i	0.227**	-0.227**
	(0.104)	(0.104)
ave_c~xt_i	-0.032	0.032
	(0.152)	(0.152)
ave_craf~i	0.141	-0.141
	(0.126)	(0.126)
ave_crea~i	0.535**	-0.535**
	(0.105)	(0.105)
ave_expe~i	0.137	0.137
	(0.182)	(0.182)
ave_crit~i	0.361**	-0.361**
	(0.171)	(0.171)
ave_comm~i	0.048	0.048
	(0.136)	(0.136)
ave_fit_i	0.431**	-0.431**
	(0.131)	(0.131)
ave_coll~i	0.640**	-0.640**
	(0.276)	(0.276)
av~ction_i	-0.055	0.055
	(0.086)	(0.086)
ave_pers~i	0.226	-0.226
	(0.145)	(0.145)
Res_Final	0.840**	-0.840**
	(0.096)	(0.096)
N	355	355
Mills ratio		
Lambda	0.387	-0.395
Rho	0.635	-0.463
Sigma	0.609	0.853

Standard errors in parentheses
* p<0.10, ** p<0.05

In our sample we have 355 observations: 157 accepted and 198 rejected applicants. As the selection equation shows, we find that Dutch applicants have higher probability of admission, and being interviewed a second time has a negative effect. There are no differences in the probability of admission among the five disciplines. Our results show that six criteria have statistically significant effects on admission, very much in line with the results reported in Table 2.5. Criteria concerned with creative and technical processes, i.e., content and creativity, receiving a high score on individual characteristics, i.e., ability to handle critique and jury member(s) wanting to professionally collaborate with the applicant, and scoring well on the fit with RABK goals criteria and the jury's final assessment are significant predictors of acceptance. In contrast with Table 2.5, we now have measured Final assessment as a residual after a regression on the 11 separate criteria. The final assessment emerges in the final jury meeting, and obviously reflects the assessments on the separate criteria. By taking the residual we use the additional information that has not already been accounted for in the separate criteria. In this simultaneous modeling of admission and performance, the letter of reference no longer has significant effect on the admission decision.

The selection equation does not support the stated prior weights of the criteria on the final decision. As Appendix A indicates, content should be the most important variable, followed by 4 equally weighted criteria (context, critique, communication and fit), next creativity and then craft and expertise. In the actual decision, collaboration and creativity have the highest weights and are statistically significant, with critique and fit in second place. Content, supposed to have the highest weight takes only third place. Clearly the jury is not doing what it thought it was doing, which is no doubt a consequence of the unstructured nature of the selection process. The conclusion is supported by the large and highly significant effect of the residual final assessment: apparently, factors not accounted for in the weighted criteria schedule have a strong influence[28].

In Table 2.6.1 we use ArtFacts.Net rankings as a measure of performance, which is the mean of the 2013 and 2015 rankings if both positive, or one year ranking if only one positive. First, in the outcome equation for the accepted applicants, our results show that receiving a high score on the fit with RABK goals criteria and the jury's final assessment are strong predictors of artists having lower (better) reputational rankings. In the outcome function for rejected applicants, content and the final jury assessment are strong predictors of lower (better) ArtFacts.Net scores. We also find that being interviewed twice (or more) in different years predicts a higher (worse) ArtFacts.Net score for the rejected, while this effect is not statistically significant for the accepted applicants. Painters, from both the accepted and rejected categories, have higher (worse) reputational rankings.

28 With the three aggregate variables in the selection equation, the residual and art are highly significant at roughly equal weights; fit with RABK goals has slightly lower weight, personality is not significant.

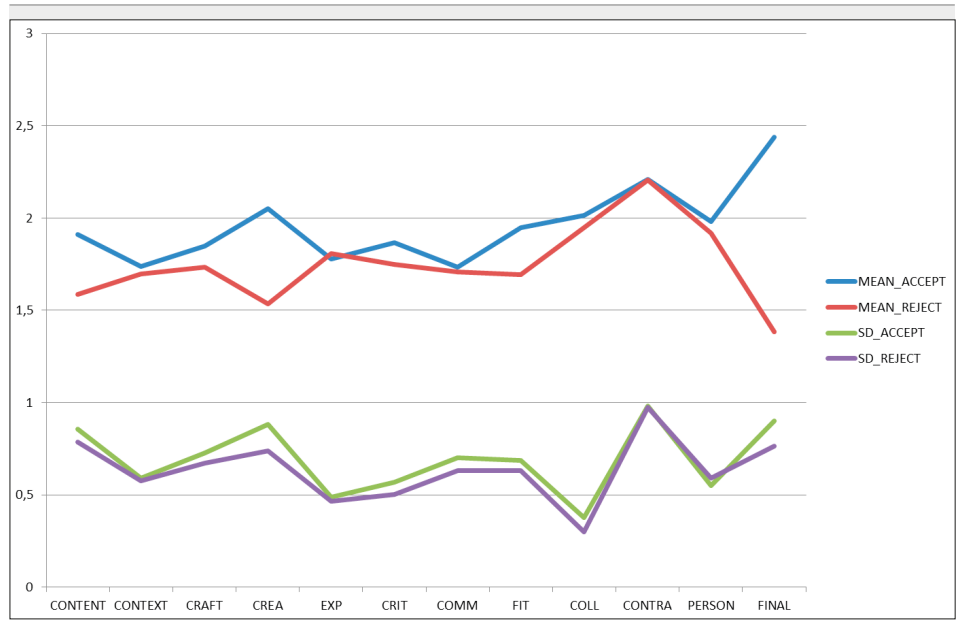
In Table 2.6.2 we measure performance by taking the log value of cumulative career sales at public auctions. With the same selection equation as before, the selectivity correction is not significant. Many of the significant variables for accepted applicants in the performance equation for reputational rankings are also significant for the accepted in the sales equation; however, these results should be interpreted with caution, as only 90 artists have non-zero sales value. Our results show that receiving a high score on criteria concerned with creative processes, i.e., a high score on the creativity criterion, predict higher sales for accepted applicants. Receiving a high score on final assessment has no effect on sales performance. In the outcome function for rejected applicants, only fit and collaboration criteria have a significant and positive effect on sales. We find that being interviewed twice (or more) in different years no effect on the performance outcome of the accepted and rejected applicants.

Formal statistical testing indicates that the performance or outcome equations for the accepted and the rejected applicants are not significantly different. We tested equality of all regression coefficients simultaneously, both including and excluding the intercepts and could not reject equality. We also tested coefficients separately, maintaining the other coefficients at their values estimated for accepted and rejected. Only two coefficients differ significantly (at 10%) for the ranking performance (communication and expertise), no difference is significant for auction sales.

The equality of the performance equations for the accepted and the rejected is a highly relevant finding of our analysis. This points to a situation where there is only one "production function" for success. All the jury can do is selection of applicants on the basis of the inputs, the variables "producing" the performance: pick the applicants with the highest values of the predictors (properly weighted). One could also imagine the situation where the production function is different in the RABK, as a more effective producer of output (performance) from given inputs. This, after all, is the goal of the RABK residency program; help promising artist to develop their talent and, presumably, to become a successful artist. If their qualifications, as measured by their criteria scores, mature in the outside world just as well as in the RABK residency, the program does not add anything, except what is going on through the selectivity correction: the qualifications that we do not observe. Indeed, the Mills ratio or lambda is significant for both the accepted and the rejected in the case of rankings, and not significant for the accepted and the rejected in the case of sales. Accepted artists have unobserved qualities that lead to higher rankings, and rejected artists have unobserved qualities that lead to lower rankings. This can be taken as a credit for the selection process: the jury recognizes more than we can observe. Of course we should not forget that our entire analysis is based on the sample of artists who passed all the initial selections: our findings only relate to those artists that have made it to the final round. One can easily anticipate substantially different results for the entire sample of applicants.

With minor and statistically insignificant differences in performance equations, the major difference among the accepted and the rejected is the difference in qualifications at entry, as a consequence of selection. Figure 2.5 indicates that the accepted score higher on all jury criteria. Most differences are modest, the highest values obtain for creativity and residual final assessment. Standard deviations are very close for the accepted and the rejected.

FIGURE 2.5 Scores on selection criteria for accepted and rejected applicants (means, standard deviations)



2.7 ROBUSTNESS TESTS

We conducted a number of robustness checks (available upon request). In our previous analysis, we used switching regressions based on parametric two-step selection regressions (Table 2.6.1 and Table 2.6.2), in which we assigned individuals who were not ranked by Artfacts.Net in either 2013 or 2015 a score of 25 (log), where the best score was the 5.01 (log) and the worst 11.82 (log). In the first robustness test, we use similar parametric selection specifications, however, we impute 15 (log) instead 25(log) for unranked artists. As earlier, we include a dummy variable coded 1 if an artist has never been ranked, and otherwise 0. These results are robust to our original specifications. In the second robustness test, we kept the unranked artists at 15(log) but included the three aggregate variables called: artistic quality, fit to the RABK goals and personality. These results are also similar to our previous findings.

In the third robustness check, we estimate two regressions for non-linear distributions of count dependent variables: Poission and probit. In the Poisson regression we use the count of auction sales as the dependent variable, and in the probit regression positive career sales are coded as 1, no career sale zero. The results from both the poison and probit regressions are comparable to the results from the performance equation for auction sales in our parametric selection models.

Lastly, we used three new operationalizations of our selection criteria variables. We apply these three alternatives variables in switching regressions to predict reputational rankings and auction sales, imputing both 25 (log) and 15 (log) for individuals who did not receive a ranking in either 2013 or 2015. First, we create a dummy variable indicating if the jury commented on a particular criterion during the evaluation of the applicant; if yes, a value of 1 was given, otherwise a 0. Second, we created a variable indicating the maximum score the applicant received per criterion, as determined by the highest score given by any of the three raters. And third, we used only the scores of one rater (one of the authors). All alternative operationalizations in all models were similar to our focal equations in the parametric switching regressions.

2.7.1 Alternative selection simulations

For the observed accepted applicants and the observed rejected applicants we have mean characteristics and mean performance outcomes. We now simulate a change in the composition of the accepted and rejected applicants by changing the admission rules, while still maintaining the proportion of applicants admitted. We do this to see if selection "by flipping a coin" or using alternative selection rules result in different performance outcomes compared to our observed sample population. First, using the entire sample (n=697), we simulate selection based upon a coin toss. Second, we replace the estimated coefficients in our main selection equation by the selection criteria weights, as rated by the selection committee members and the head of the residency program, for a sub-sample of applicants with jury notes (n=355), while keeping all other coefficients at their estimated value. We then select the top 43% of these applicants, which is proportionally equivalent to the observed sample, as accepted and reject the rest. We are aware that alternative admission rules can change artists' status; namely, we may admit artists who were originally rejected and reject those who were originally admitted; hence there is a selectivity issue. However, we argue that with these data selectivity is inconsequential because the two performance equations from the observed sample population are not statistically different from one another. Therefore, we may use the observed output even if the applicant, with the alternative selection rules, has been reallocated to another status.

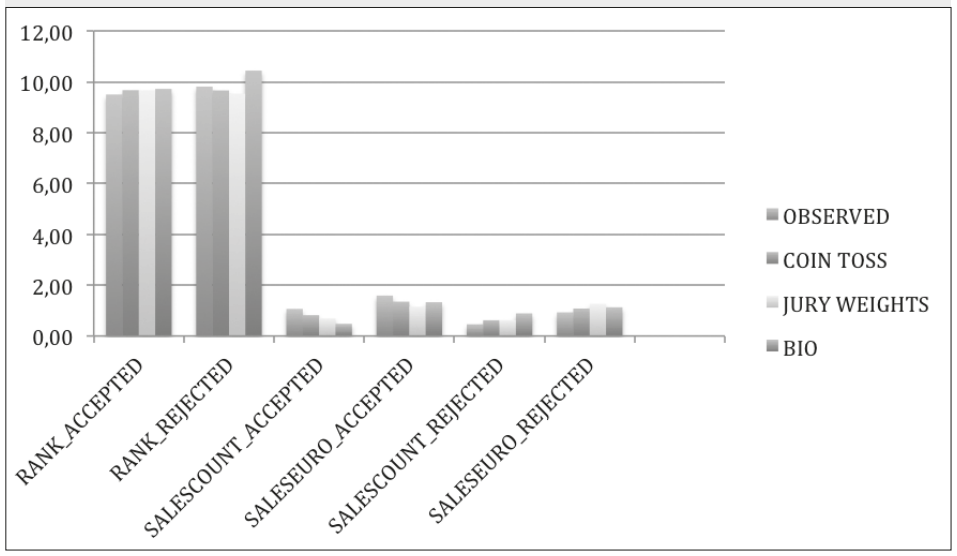
Comparing the mean performance outcomes of our observed sample population to the estimated performance outcomes from the coin toss and jury weight simulations, we observe that the accepted observed sample population performs slightly better than the simulations. For instance, the log mean artistic ranking for accepted applicants in our observed sample population is 9.50, while the log mean artistic rankings based on chance and jury weights are slightly worse, 9.67 and 9.66, respectively. The accepted observed sample population also sells a slightly greater number of artworks at auctions for a slightly higher price compared to the performance outcomes of the two simulations. The rejected artists from these simulations perform slightly better than the rejected artists from the observed sample population.

Based upon this comparison, it is possible to speculate as to why the accepted observed sample population performs slightly better across both performance outcomes compared to the accepted artists in the simulations, and why the rejected in the observed sample population perform slightly worse across both performance outcomes compared to those rejected in the simulations. The coin-toss and jury weight simulations include artists who were accepted according to alternative selection rules. This means that some applicants were accepted in our observed sample population yet rejected in the simulations, and visa versa. For instance, of the 157 artists accepted according to the jury weight simulation, 86 were also accepted in the observed sample population, while 71 of these "newly" accepted artists were rejected by RABK. The coin toss simulation consisted of 299 "newly" accepted artists, of which 114 were also accepted in the observed sample population, while 185 were rejected.

One possible way to interpret these differences in performance outcomes between the observed sample population and the simulations is that a residency at RABK is a signal of quality to the international contemporary art market. Such a signal of quality is important, since, as noted before, there are no objective measures to determine the quality of artists or artworks and the market in general is also characterized by high uncertainty (Caves, 2000; DiMaggio, 1987; Yogev, 2010; Velthuis, 2013). Signaling theory (Spence, 1973) describes how signals, e.g. an elite education or participation in a prestigious program, help create a separating equilibrium based on an inverse cost quality relationship, in which receiving or sending a signal is less costly for a high quality producer compared to those of lesser quality (Spence, 1973; Connelly, Certo, Ireland and Reutzel, 2011). It is possible that RABK acts as a proxy of quality in the international contemporary art market, by reducing uncertainty about quality caused by incomplete or imperfect information. Precisely the fact that the RABK can be seen to represent peers can increase the credibility of the signals it produces to other peers (Ebbers and Wijnberg, 2012) and experts who tend to respect peer judgments in high art. This would help explain the slightly better performance outcomes of the accepted applicants, compared to the simulations, and the slightly worse performance outcomes of the rejected applicants, compared to the simulations.

Another possibility is that RABK functions as it aims: namely, to help develop talent in visual arts. Perhaps the accepted applicants learn something to help them develop and deepen their art practice and/or are granted access to an international network of art critics, curators and collectors. Or perhaps the selection committee systematically picks excellent artists. One thing is for sure, changing the selection rules to fit the weights of the jury (as described by committee members and the head of residency) does not increase applicants' performance outcomes, nor does simply chance. Apparently, being accepted to RABK positively affects artists' future performance.

FIGURE 2.6 Performance outcomes (log) for the observed sample population and simulations



2.8 CONCLUSIONS

The Rijksakademie van Beeldende Kunsten (RABK) has kept detailed records of the selection processes for its internationally highly prestigious program that provides promising artists a two-year residency with full facilities. We combined information from this selection process with data on later performance of accepted and rejected artists, to assess the selection process in terms of artistic prestige and economic success. Almost 75% of both accepted and rejected artists in our analysis had rankings of artistic standing, suggesting that most remained active in the art market, although only a small percent (13%) ever sold artworks at an art auction. This can be explained by the fact that artists active on the primary art market usually sell their artworks through art galleries or directly to art collectors, rather than at art auctions (Prinz, Piening and Ehrmann, 2015). Nevertheless, a skewed distribution of success is common in the cultural industries (Caves, 2000), as can be observed with our data, as those artists who sold artworks at art auctions career average sales totaling €167,552.10 (minimum €228.78 and maximum €7,762,842).

As mentioned earlier, RABK aims to admit exceptionally talented artists and to facilitate their development to the top of the international art world. We analyze the interview round of the admission procedure where a selection committee selects applicants in an informal process relying on objective biographical data and subjective assessments of applicants' artistic qualities and potential for development. These assessments are based on criteria that have not been formalized and weighted in explicit statements but were retrieved through non-participatory observation and interviews with selection committee members. The discussion on criteria scores for applicants is not structured along formal rules; to distill this information three raters interpreted the comprehensive jury deliberation notes and assigned scores. Our first conclusion is that the selection committee operates differently than the way she thinks she does: the weights of the criteria that we estimate from the observed admission decisions are not equal to the weights intended by the selection committee.

Nevertheless, in a general sense, the selection committee seems to have realized their aim of selecting exceptional talent and facilitating development to leading positions in the art world. The admission decision appears to relate positively to the applicant's later success; artists with better artistic ranking and more sales have higher probability of admission. The mean artistic ranking of accepted applicants is indeed (weakly) statistically higher than the mean ranking of rejected applicants; visual inspection indicates that accepted applicants have a smaller lower tail in their frequency distribution of rankings than the rejected applicants.

Analysis of the admission process shows that among the biographic data, nationality, a reference letter and repeated participation in the jury process have a significant effect on the admission decision. Among the separate jury criteria, significant monotonic effects are not the rule. Scoring above average on content, creativity, collaboration and fit have a clear positive effect, scoring below average on these same criteria has a clear negative effect.

The final jury assessment, combining much information collected in the selection processes, on its own has a strong effect on the admission decision. Averaging the assessments into three variables shows that artistic quality and fit to the RABK goals have significant effect on the admission decision, while personality does not. Our second conclusion is that multiple selection criteria predict admission, although the overall final jury assessment seems to have a stronger effect.

From a switching regression model, where the effect that unobserved qualities governing admission may have on performance is taken into account, these unobservables appear to increase the likelihood of admission and the likelihood of superior performance for the accepted applicants, while reducing performance for the rejected. Thus, admission increases the difference in performance outcomes of the accepted and the rejected and as such may be inferred to function as a signal of quality; although the performance equations for accepted and rejected artists are not statistically different from each other. This means that the effect of measured qualities as perceived by the jury on later performance does not differ among applicants who have resided at RABK for two years and applicants that have not been admitted. The result is more positive however, on the joint effect of variables that we, as outsiders, do not observe but the selection committee does. The Mills ratio is significant for both the accepted and the rejected in the case of rankings, but not significant for sales. Accepted artists have unobserved qualities that lead to higher rankings and rejected artists have unobserved qualities that lead to lower rankings. This can be taken as a credit for the selection process: the jury recognizes more than we can observe. This also may explain why the jury's final assessment of an applicant, purged from the effect of the specified selection criteria, has a significant effect on the admission decision. Yet, this final assessment score only has substantial and significant effect on artistic ranking, and no effect on sales performance. Our third conclusion is that multiple selection criteria, and the final jury assessment, predict artistic rankings but not sales. Artistic ranking is predominately determined by peer selection. Perhaps the selection committee, consisting primarily of peers, recognizes and acts upon the shared beliefs and agreements about novelty, contribution and maintaining high quality standards shared by others peers across the board (Kostoff, 1997; Eisenhart, 2002), which would help explain the effect on rankings but not on auction sales, which might be less strongly influenced by signals of peer appreciation and more by signals coming from experts or even other customers (Ebbers and Wijnberg, 2012; Gemser, Leenders and Wijnberg, 2008).

In the end we must conclude that our findings are very much in line with the literature. Predicting success in creative occupations, in other words picking the winners, is quite a challenge. Unstructured interviews have a poor record. Careers of accepted and rejected artists are different because artists unobserved attributes, such as ability and perseverance, may differ at entry, and not simply because the residency substantially increases the pay-off to their qualities. The caveat is of course that our entire analysis is based on the sample of artists who passed all the initial selections: our findings only relate to those artists that have made it to the final round.

2.9 APPENDIX

APPENDIX A: Selection criteria per round

Selection round	Interview round	Criteria	Description	Importance
15%	*15%*	*Socio-Demographic*		
			Candidate is in target age group (26-36)	4
			Candidate is male/female	1
			Candidate is male/female (Interview)	4
			Candidate has independent art practice > 2 years	4
			Country of origin (Silent Zones)	4
			Country/city of current residency	4
10%	*0%*	*Education*		
			Previous art education (bachelor/master degree)	3
			Previous art residencies	4
			Previous higher education outside art	2
			Graduating with honors	2
10%	*0%*	*Number of Times Applied*		
			Candidate applied previously	1
			Amount of times previously applied	2
			Highest selection previously achieved (Interview)	4
			Candidate knows RABK advisor(s)	4
			Candidate knows RABK past /present resident(s)	4
15%	*0%*	*Letter of Recommendation*		
			From RABK advisor	4
			From well-known artist	4
			From unknown artist	2/3
			From prominent non-artist	3
5%	*0%*	*Written application*		
			Text about work/expectations from RABK	2/3
45%	*80%*	*Art work/ Portfolio*		
(25%)	(20%)	Content	Research based/ work in progress	4
			Evolution of work over time/ continuity	4
			Problem finding orientation	4
(20%)	(15%)	Context	Art historical reference(s)	4
(10%)	(5%)	Craft	Technical aspects	2
(15%)	(10%)	Creativity	Innovative ideas/ methods	5
			Originality	4
			Novel contribution	5
(0%)	(5%)	Expertise	Diverse expertise (art/non-art)	4
(10%)	(15%)	Critique	Open to assessments/criticism/ change	4
(10%)	(15%)	Communication	Purpose/reason for making work	5
			Professional awareness/confidence	4
(10%)	(15%)	Career Fit		5
		Personal[29]		
		Collaboration	Jury member would like to work with artist	
		Contradiction	Jury had opposing assessments	
		Personality	Jury liked the artist	
		Final assessment		

29 Personal characteristics or final assessment did not receive weights; these criteria were recorded after analysis of the jury notes.

APPENDIX B: Applicants by cohort year

RABK : All Applicants during observation period

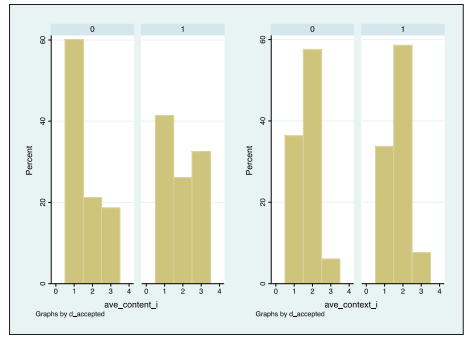
Year Applied	Year Started	Applications	Selection	Interview	Accepted	Rejected
1995	1996	680	680	69	33	36
1997	1998	672	672	71	25	46
1999	2000	666	666	51	27	24
2000	2001	606	606	65	32	33
2001	2002	634	634	56	27	29
2002	2003	950	950	54	29	25
2003	2004	947	670	59	27	32
2004	2005	844	469	62	23	39
2005	2006	830	440	64	25	39
2006	2007	1096	565	75	25	50
2007	2008	1324	296	71	26	45
Total:	*11 years*	*9249*	*6648*	*697*	*299*	*398*

APPENDIX C: Variables

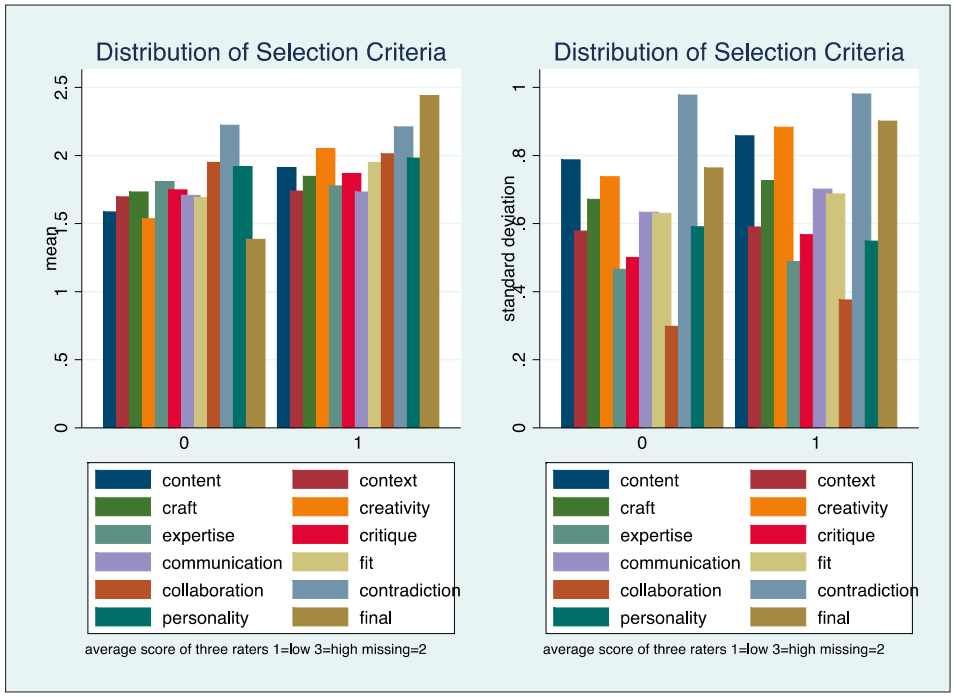
Variable name:	Description:	Variable type
d_accepted	accepted to RABK(1), rejected from RABK(0)	binary
d_female	gender: female (1), male (0)	binary
age_interview	age at interview	continuous
age2	age at interview (squared)	continuous
career_age	year applied minus 2015	continuous
career2	year applied minus 2015 (squared)	continuous
d_nation	Dutch nationality (1), other (0)	binary
d_no_art_ed	no formal previous art education (1), formal previous art education (0)	binary
d_ref_letter	reference letter (1), no reference letter (0)	binary
d2xinterview	went to interview round in subsequent year(s) after being rejected (1), interviewed once (0)	binary
d_sculpture	main artistic discipline sculpture (1), otherwise (0)	binary
d_video_film	main artistic discipline video/film (1), otherwise (0)	binary
d_drawing	main artistic discipline drawing (1), otherwise (0)	binary
d_photography	main artistic discipline photography (1), otherwise (0)	binary
dauctionc	auction sales (1), no auction sales (0)	binary
dArtFacts.Net2013	2013 ArtFacts.Net rank (1), no rank in 2013 (0)	binary
dArtFacts.Net2015	2015 ArtFacts.Net rank (1), no rank in 2013 (0)	binary
d_2015_13	2015 ArtFacts.Net rank in either 2013 or 2015 (1), otherwise (0)	binary
d_jury1995_1998	change in selection process: jury specialized in core discipline of the applicant made the final selection (1), 1999-end of observation period multi-disciplinary jury assessed the work from all applicants (0)	binary
content	development of the artist/artworks over time. low score (1), missing score (2), high score (3)	discrete
context	artists' understanding of art historical references. low score (1), missing score (2), high score (3)	discrete

craft	technical aspects of the artwork/ ability in actually making the artwork. low score (1), missing score (2), high score (3)	discrete
creativity	innovative aspects/ novel contribution of the artwork. low score (1), missing score (2), high score (3)	discrete
expertise	art education or combinations with other academic disciplines low score (1), missing score (2), high score (3)	discrete
critique	openness of the artist is to criticism and change low score (1), missing score (2), high score (3)	discrete
communication	ability to communicate about their artworks and reason for making the artworks low score (1), missing score (2), high score (3)	discrete
fit	artists' career and the fit within the RABK community low score (1), missing score (2), high score (3)	discrete
collaboration	would the jury member(s) like to have the artist as a colleague low score (1), missing score (2), high score (3)	discrete
contradiction	level of opposing views between jury members about an applicant: complete or partial disagreement between jury members (1), missing score (2), complete agreement (3)	discrete
personality	does the jury liked the personality of the applicant low score (1), missing score (2), high score (3)	discrete
final assess	jury final assessment of the applicant overall negative evaluation (1), overall positive evaluation (2)	discrete
art	average score of the selection criteria: *content, craft, creativity* Low score (1), no score (2), high score (3)	discrete
rabk_goals	average score of the selection criteria: *fit, context, critique* Low score (1), no score (2), high score (3)	discrete
personality	average score of the selection criteria: *Expertise, Communication, Collaboration, Contradiction, Personality* Low score (1), no score (2), high score (3)	discrete
log2013ArtFacts.Net	ArtFacts.Net score in 2013 (log)	continuous
log2015ArtFacts.Net	ArtFacts.Net score in 2015 (log)	continuous
logmean2013_2015	average 2013 and 2015 ArtFacts.Net scores. For those ranked in either 2013 or 2015, the average rank is the rank they received in that particular year.	continuous
res_log2015ArtFacts.Net	residual of log2015ArtFacts.Net purged of career age and career age2 effects	continuous
res_log2013ArtFacts.Net	residual of log2013ArtFacts.Net purged career age and career age2 effects	continuous
res_logmean2013_2015	residual of logmean2013_2015 purged career age and career age2 effects	continuous

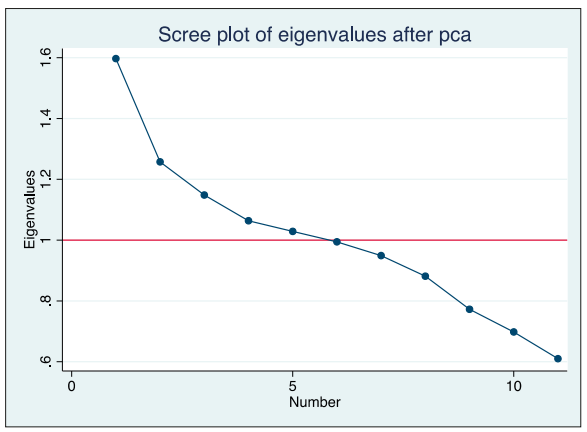
APPENDIX D: Distributions of multiple selection criteria score among the sample



APPENDIX E: Means and standard deviations of selection criteria



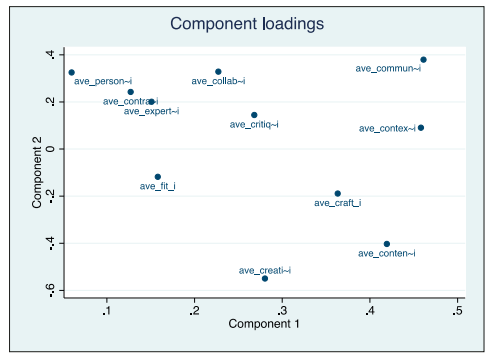
APPENDIX F: Principle components analysis (PCA)



Principal components/correlation — *Number of obs = 355*
Number of comp. = 3
Trace = 11
Rotation: orthogonal varimax (Kaiser off) — *Rho = 0.3639*

Rotated components (blanks are abs(loading) <.3)

VARIABLE	COMP1	COMP2	COMP3	UNEXPLAINED
CONTENT		0.5332		.5149
CONTEXT	0.4665			.6335
CRAFT	0.3335		-0.4682	.5209
CREATIVITY		0.6910		.3679
EXPERTISE				.865
CRITIQUE			0.5307	.5597
COMMUNICATION	0.5713			.479
FIT				.8925
COLLABORAION	0.3621			.7757
CONTRADICTION				.8677
PERSONALITY			0.6335	.5203



The Kaiser-Meyer-Olkin measure of sampling adequacy shows that PCA is acceptable at kmo >.50; our overall score was 0.5209, with the lowest value being for the collaboration variable which had a value of 0.4432.

We also conducted a factor analysis to check for underlying lower dimensionality of the data set. The three factors in our factor analysis were similar to the three components in our PCA.

APPENDIX G: Correlation Matrix

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.
1. AVE_CONTEN~I	1.0000														
2. AVE_CONTEX~I	0.0912	1.0000													
3. AVE_CRAFT_I	0.1611	0.1535	1.0000												
4. AVE_CREATI~I	0.2581	0.0695	0.0565	1.0000											
5. AVE_EXPERT~I	0.0446	0.0737	-0.0156	-0.0019	1.0000										
6. AVE_CRITIQ~I	0.0051	0.0708	0.0121	0.1204	-0.0401	1.0000									
7. AVE_COMMUN~I	0.0815	0.2158	0.1423	-0.0235	0.1024	0.1277	1.0000								
8. AVE_FIT_I	0.0678	0.0826	-0.0241	0.0535	-0.0373	0.0254	0.0419	1.0000							
9. AVE_COLLAB~I	0.1095	0.0394	0.0032	-0.0888	0.0594	0.1168	0.1110	-0.0070	1.0000						
10. AVE_CONTRA~I	-0.0249	0.0640	0.0485	-0.0198	-0.0071	-0.0087	0.0683	0.0258	0.0494	1.0000					
11. AVE_PERSON~I	-0.0067	0.0050	-0.1070	-0.0029	0.0113	0.0851	0.1164	-0.0051	-0.0210	0.0259	1.0000				
12. AVE_FINALJ~I	0.2585	0.1185	0.1510	0.2273	0.0128	0.1320	0.0786	0.1890	0.0930	0.1137	0.0160	1.0000			
13. ART QUALITY	0.7345	0.1526	0.5551	0.6932	0.0156	0.0722	0.0930	0.0536	0.0117	-0.0023	-0.0522	0.3235	1.0000		
14. RABK GOALS	0.0924	0.6179	0.0730	0.1285	-0.0030	0.5412	0.2029	0.6665	0.0737	0.0456	0.0411	0.2429	0.1492	1.0000	
15. IND QUALITY	0.0566	0.1628	0.0461	-0.0427	0.3602	0.0914	0.5663	0.0205	0.3047	0.6725	0.4331	0.1342	0.0273	0.1438	1.0