

Journal Pre-proofs

Factors explaining differences in wine experts' ratings: The case of gender, credentials, occupation and peer effects

Florine Livat, Hervé Remaud, Marta Fernández-Olmos

PII: S0963-9969(24)00759-2
DOI: <https://doi.org/10.1016/j.foodres.2024.114689>
Reference: FRIN 114689

To appear in: *Food Research International*

Received Date: 1 February 2024
Revised Date: 17 June 2024
Accepted Date: 26 June 2024

Please cite this article as: Livat, F., Remaud, H., Fernández-Olmos, M., Factors explaining differences in wine experts' ratings: The case of gender, credentials, occupation and peer effects, *Food Research International* (2024), doi: <https://doi.org/10.1016/j.foodres.2024.114689>

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2024 Published by Elsevier Ltd.



Factors explaining differences in wine experts' ratings: the case of gender, credentials, occupation and peer effects

Florine Livat^a, Hervé Remaud^b and Marta Fernández-Olmos^c

^a Kedge Business School, 680 cours de la Libération, 33400 Talence, France. florine.livat@kedgebs.com (corresponding author)

^b Kedge Business School, 680 cours de la Libération, 33400 Talence, France. herve.remaud@kedgebs.com

^c University of Zaragoza, Pedro Cerbuna 12, 50009 Zaragoza, Spain. maferno@unizar.es

Abstract

The aim of this article is to better understand why experts give different scores to the same wines in identical tasting environments. This research focuses on the personal characteristics of experts (or judges), such as their gender, industry credentials, and occupation within the wine industry, and examines how judges respond to their peers' characteristics. Using a dataset of 5,395 wines judged in the 2022 International Wine and Spirits Competition, we analyse 18,224 scores from different judges. We estimate a series of grade equations at the judge level to understand why a same wine received different scores from judges. A first model makes use of the panel structure of the dataset, incorporates wine fixed effects, and focuses on the personal characteristics of judges. A second model encompasses characteristics of the judging team, without the inclusion of wine fixed effects. At large, on-trade buyers give lower scores than off-trade buyers, as well as female judges compared to male ones. While credentials are not a very significant factor *per se*, they do have the potential to generate peer effects. Judges tend to be more generous in their assessments when they are assigned to a team with Master of Wine judges. Conversely, they are also consistently more severe when the number of female judges on the team increases. Estimation results converge across sub-datasets, with the exception of sparkling wines. Given the feminisation of the wine industry, in terms of producers, consumers, and experts, the severity of female judges could be beneficial in terms of social welfare.

Key words: wine experts, ratings, gender, credential, occupation, peer effects.

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Journal Pre-proofs

“[D]ifferences amongst tasters, reflecting each individual’s physiology, experience and knowledge, are valid data in themselves rather than ‘error in the machine’.”

Parr (2019, p. 230)

1. Introduction

Wine expertise is not always characterised by consensus. On the contrary, it is marked by a lack of consistency between experts (Ashton, 2012, 2017). For instance, different experts may perceive and understand the elegance of a wine in very different ways (Shepherd et al., 2023). It brings with it a number of challenges, including for the consumer in terms of knowing which expert to trust and take into account when making a decision, and for the producer and seller in terms of knowing which competition to enter and which expert opinion to promote or disseminate. Wine competitions may also wonder what profile of judge they should welcome. The aim of this article is to improve understanding of the reasons why experts may give different scores to the same wines assessed in identical tasting environments.

Wine experts have extensive knowledge about wine (Grohmann et al., 2018). They improve and develop their skills through deliberate practice and wine tasting (Croijmans and Majid, 2016). When evaluating wines, experts describe them using numerical ratings and/or tasting notes. As such, experts play an important role in many markets, shaping consumer judgements and purchasing decisions by mitigating the asymmetric information problems inherent in experience goods such as wine. Expert opinion is positively correlated with consumer responses, including wine evaluation (Chocarro and Cortiñas, 2013), willingness to pay (Weerasekara and Streletskaya, 2024), and the demand for wine (Hilger et al., 2011), with different effects depending on wine price (Bonnet et al., 2020) and if supplemented by community reviews (Anderson and Magruder, 2012). However, welfare is at risk here because intermediaries providing information can act strategically and influence the market outcome (Lizzeri, 1999). Consumers may suffer welfare losses if expert opinion is biased. This is the case, for example, if experts' liking ratings vary systematically when the origin of the wine is disclosed, reducing international trade in some particular wines and limiting the range of food and wine pairings on offer (Depetris-Chauvin et al., 2024). The way experts perceive and interpret wine can also be influenced and shaped by the conversation they have with wine producers, especially if those producers have achieved a high status in the wine hierarchy, where storytelling, more than wine quality, is part of that conversation (Humphreys and Carpenter, 2018). Nevertheless, Cardebat and Livat (2016, p. 54) note that “the welfare effect of systematic variations in the appraisal of different experts depends on how consumers use expert opinions” and add that “a greater variety of opinions across experts should enable consumers to find an expert whose tastes are close to theirs, leading to higher levels of consumer welfare.” Consequently, if discrepancies in expert ratings are primarily a *de facto* expression of personal preferences rather than a manifestation of bias, consumers may be able to identify an expert whose preferences closely match their own. Of course, not all consumers will engage in a process of finding an expert with similar tastes, due to the time, cognitive effort and resources required, but many would, regardless of the positioning and price range of the wines being rated, as expert ratings are not just for premium wines.

Wine tasting involves sensory analysis and complex cognitive processes of perception, conceptualisation, memory activation, imagination and communication (Parr, 2019). These processes are specific to each expert, depending on their own training, experience, culture (Parr, 2019) and even childhood olfactory experiences (Chu & Downes, 2000). In other words, perceptual activity is biased because who we are and what we already know and have

experienced influences what we perceive. More generally, among other factors, autobiographical memory plays an important role in sensory experience. As noted by Lesschaeve (2007), there are different types of experts with different training, which include wine makers, wine sellers and wine writers. We expect that different wine experts following a similar process (i.e. visual, olfactory and gustatory examination) and evaluating the same wine in the same environment will not necessarily reach the same rating.

As sensory analysis is a subjective experience, social influences have often been excluded from the scope of researchers, particularly in experimental settings where subjects are seated in individual booths and social interaction is deliberately limited (see Sáenz-Navajas et al. (2016) and Nunes et al. (2017) for experiments with wine experts). However, sensorial experiences can be subject to social influences. For instance, the commensality dimension of wine consumption can imply a restaurant or wine bar venue (Livat et al., 2024), or simply gathering family or friends (Jaud et al., 2023). Wine competitions frequently rely on the assessment of a panel of judges, who are tasked with evaluating the same wine (Berg et al., 2022). The phenomenon of peer effects among judges has been documented in courtrooms where a third party is judged (Eren and Mocan, 2020). However, there is a gap in the literature regarding the existence of peer effects among experts in the case of experience goods, particularly those that associate taste and hedonic characteristics like wine.

We contribute to the empirical literature on wine expertise by further analysing differences in expert ratings, examining the effect of individual characteristics as well as peer or team effects. Existing research is based on experimental data, which allows for a controlled setting, but with specific wines chosen for the experiment, sometimes not aligned with the markets. Furthermore, the substantial amount of original data analysed in this study is based on a large and diverse group of 91 wine experts who provided individual grades. This allows us to investigate peer or team effects. Using an original dataset of 5,395 wines provided by a UK-based awarding institution, we estimate a series of grade equations and focus on personal characteristics of the experts, such as gender, industry credentials and occupation within the wine industry, and on peer or team effects. To the best of our knowledge, these personal characteristics of wine experts have not been considered in combination, nor have potential peer effects been examined.

The remainder of our article is structured as follows: Section 2 presents background knowledge on the differences between expert ratings, Section 3 is dedicated to the materials used to analyse the wine ratings, Section 4 presents the econometric analysis of the wine ratings, and Section 5 discusses the results and provides a conclusion.

2. Differences between experts' ratings

Previous empirical literature has extensively analysed expert scores and multidimensional scorecards (e.g. combining acidity, balance, finish, etc.) using a variety of tools ranging from descriptive statistics (Cliff and King, 1996) to complex econometric methods (Paroissien and Visser, 2020). These analyses often focus on objective characteristics of the wine, such as colour, region of origin and vintage, as well as organoleptic characteristics of the wine (Combris et al., 1997; Grohmann et al., 2018; Parr et al., 2007). Experts habitually converge in their judgments of overall wine quality, but Gawel and Godden (2008) note that they may weight specific aspects of quality differently. It is therefore possible to classify and cluster experts with

similar styles according to how they rate individual components of quality (Cliff and King, 1996; Scaman et al., 2007).

These approaches suggest that experts have different tastes and therefore may not give the same score and comments to a given wine. On the contrary, some prefer specific wine styles (Cardebat and Livat, 2016), while others are more generous with wines they have never encountered and for which they have no reference point (Livat and Remaud, 2021). Indeed, these are mental representations that can vary according to the experience, practice and background of wine experts (Otheguy et al., 2021). Another line of research has analysed the influence of the tasting environment and context, such as the type of tasting glass (Fischer and Loewe-Stanienda, 2000; Spence, 2020), the colour lighting of the tasting room (Spence et al., 2014), the colouring and visual style of floral arrangement laying the table (Rodrigues et al., 2023), the evaluation condition, such as the type of sensory analysis (Sáenz-Navajas et al., 2016), or the scoring system, namely above 20 or 100 points, which does not explain the differences between judges (Parr et al., 2006). Finally, like consumers, wine experts can be influenced by extrinsic cues such as country of origin (Depetris Chauvin et al., 2024) or brand information (D'Alessandro and Pecotich, 2013). More recent research has looked at the characteristics of the judges themselves.

2.1 Culture

Cognitive processes, including wine appreciation, vary across cultures, often defined in terms of geographical location (see Rodrigues and Parr, 2019, for a review). Here, if most existing studies have focused on consumers, some have also considered experts or professional wine tasters. For example, Sáenz-Navajas et al. (2013) show that the perception of quality does not change between French and Spanish oenologists. Similarly, Valentin et al. (2016) find that the perceived quality of Pinot noir wines does not differ between French and New Zealand wine experts. Therefore, the culture of origin, linked with repeated exposure to specific wines, does not appear to be the main factor affecting quality perceptions (Sáenz-Navajas et al., 2013), but Suarez et al. (2023) note that cultural differences should be more pronounced for wines with different sensory profiles. They show that the cultural background of wine experts, combined with their familiarity with the product, influences the perception of quality.

2.2 Expert background and occupation

Different profiles of judges exist, with different types of background and training (Lesschaeve, 2007) and different professional objectives (Shepherd et al., 2023; Honoré-Chedozeau et al., 2024). For example, winemakers, often with a background in chemistry, are trained to detect defects, adopt a quality control perspective and are oriented towards consistency (Joy et al., 2019), either at the winery level (i.e. consistency of winery style) or at the regional level (i.e. consistency of typicity). Other wine experts, such as sommeliers, wine judges, wine journalists and wine educators, are oriented towards independent judgments of diverse wines targeting consumers (Grohmann et al., 2018). Similar results are obtained for tasting narratives, where winemakers use technical terms related to the presence or absence of defaults, sellers are storytellers to make wines appealing, and critics rely on metaphors to elaborate their tasting experience (Honoré-Chedozeau et al., 2024). These differences also exist in mental representations, the core of which is shared universally by wine experts and the peripheral

elements of which are based on experts' specific backgrounds and professional concerns (Otheguy et al., 2021).

2.3 Training and credentials

The wine industry offers a range of credentials to signal wine expertise and help consumers choose an expert they can trust. For example, Orduña-Malea et al. (2021) note that the Master of Wine (MW) certification is the most prestigious in the wine industry for commercial purposes. Marlowe et al. (2017) refer to several certifying bodies for wine credentials, either nationally or globally: The Court of Master Sommeliers, the Wine and Spirits Education Trust (WSET), the Society of Wine Educators, the Culinary Institute of America, the International Sommelier Guild, the Sommelier Society of America, and the International Wine Guild. These bodies offer different training programmes and develop different skills. For example, sommeliers are trained in wine service, while WSET certification ranges from basic wine tasting skills (Level 1) to a detailed understanding of viticulture and winemaking (Level 3). Grohmann et al. (2018) interpret some of the differences between wine experts in relation to these credentials, rather than in terms of the geographical location of the experts: while all experts agree on the overall quality of the wines they judge, their perceptions of subjective sensory attributes differ; in particular, experts with a sommelier background seem to be more aligned with the Old World style of winemaking, while WSET recipients are closer to the New World, where what some experts consider to be a low level of wine fault may be seen by others as simply elements of certain wine styles. In the same vein, Brochet and Dubourdieu (2001) suggest that wine professionals' training may influence the development of their own prototypes for representing wine.

2.4 Gender of experts

The impact of expert gender on wine expert ratings has not yet been fully explored, despite Cawley's (2018) suggestion that female and male experts may use different wine language, and Bodington and Malfeito-Ferreira (2018) who did not measure any significant difference among scores given by male and female wine judges. Nevertheless, empirical evidence indicates that gender differences in preferences do, in fact, exist (Falk and Hermle, 2018). Consumer-oriented research indicates that taste preferences may differ by gender (Lombardo et al., 2023). Gender differences have been extensively studied in the context of wine tasting (Pérez-Jiménez et al., 2022), wine preferences (Sena-Estevés et al., 2018), and emotional responses to wine (Mora et al., 2018). Although little research has been conducted on gender differences among wine experts, an interesting gender bias has been highlighted by Kaiman and Bru (2024), who studied the gendered phonetic properties of varietal names: if, on average, feminine names receive lower scores than masculine ones, i.e. wine ratings are characterised by a phonetic gender gap, white wines with more feminine names receive lower ratings from female experts while red wines with more masculine names receive lower ratings from male experts. Overall, gender appears to be a relevant factor in other industries, like in finance and accounting, for instance, where the gender of financial experts has been shown to affect their performance (Beckmann and Menkhoff, 2008; Zalata et al., 2018; Komal et al., 2023), or in courtrooms, where female judges are harsher than their male counterpart (Eren and Mocan, 2020).

2.5 Peer effects

The formation of opinions and judgments by experts or critics is not always an isolated process. Rather, they are influenced by the wider environment in which they operate, particularly when they are part of a panel forming a judgement during the same judging session. Peer effects are a special case among social effects, because of peers' similarity. Portyanko et al. (2023, p. 516) noted that "peers can constitute a major part of an individual's social circle as they include acquaintances". It can be reasonably assumed that a significant proportion of wine judges have had the opportunity to meet and interact with one another at international wine trade shows, fairs and competitions.

The phenomenon of peer effects is well documented among judicial judges in courthouses. (Martín-Román et al., 2015; Eren and Mocan, 2020; Holden et al., 2021). Peer effects may arise from deliberation or negotiation on the part of both parties or from prior knowledge of the preferences of the other party (Miles, 2012). Furthermore, they have been evaluated in other contexts, such as education, the workplace, and entrepreneurship. In these settings, peer effects are often driven by conformism and social pressure (Villevall, 2020), perceived status and relative influence within the group (Wilkinson and Fung, 2002), talent and high performance (Hendricks et al., 2023), or seniority (Portyanko et al., 2023). The observable characteristics of one's peers, such as gender, can also influence one's own opinion and behaviour. For example, Boyd et al. (2010) and Songer and Crews-Meyer (2000) have found that male judges exhibit different behaviours when a woman is present on a panel with them. Eren and Mocan (2020) identified that the severity of punishment increases when the proportion of female peers increases. These effects can be attributed to exogenous peer effects, or contextual effects (Manski, 1993), whereby an individual's behaviour is influenced by the fixed characteristics of the team.

3. Data

3.1 IWSC data

The dataset analysed here is provided by the International Wine and Spirits Competition (IWSC), a UK-based awarding institution with 50 years of experience and a reputation for professionalism within the industry. The IWSC organises wine tastings to recognise and reward the best products in the wine and spirits industry, with the aim of enhancing consumers' drinking experiences. As mentioned on the IWSC website, "[j]udges are selected for their knowledge and experience, and in order to ensure that products are evaluated fairly and with consumer expectations in mind, qualified and knowledgeable consumer judges work alongside [masters of wine] (MWs) and prominent trade judges" (see <https://www.iwsc.net/people/judges/wine>).

Wines are evaluated on a 100-point scale and awarded medals by 91 judges who are organized into different teams. Tasting sessions typically occur in April and May each year. The judges' teams can vary from session to session, and some judges may be asked to taste specific regions or varieties, such as champagne and sparkling wine. A chair leads each session, ensuring consistent medal scores and resolving any disagreements. IWSC ratings are based on blind tastings, guaranteeing anonymity by pre-pouring wines into numbered glasses. To avoid palate fatigue, IWSC expert judges assess approximately 65 wines per day. Wineries must pay a fee of £170 plus VAT per wine entry and ship their wines to the UK to enter the competition.

The IWSC has provided a dataset containing information on all wines judged in 2022. A total of 5,395 wines from around the world were assessed, with ratings disaggregated at the judge level. Each wine is evaluated by a unique team of two to seven judges and identified with a unique reference, resulting in 18,224 ratings. Each wine in the dataset is evaluated by all members of the same judging team, resulting in multiple appearances of the same wine in the dataset. Prior to the commencement of each tasting session, the IWSC assigns judges to a judging team. Consequently, the composition of the judging teams is not determined by the wine judges themselves, thereby eliminating the potential for self-selection bias. They discover their peers before entering the tasting room and assessing the wines: first and last names are clearly displayed, along with their Master Sommelier or Master of Wine status, if any. The composition of the teams of judges can vary due to various factors, such as health reasons or simply availability due to personal and professional constraints, and can be considered here as exogenous.

For each wine, the individual rating given by each judge and a tasting note are known, along with objective characteristics such as the product name, colour, whether it is sparkling or still, country and region of origin, vintage, level of alcohol, whether it is blended, the varieties used and their proportions, and whether it benefits from a certification of origin (e.g. AOP or IGP in France, PGI, PDO, DOC or DOCG in Italy, DOC or DOCG in Spain, AVA in the USA). These wines originate from various regions worldwide, including emerging wine regions such as Ningxia (China), Kent (England), and Baja California (Mexico). Table 1 presents the main characteristics of the wines tasted.

[Insert table 1 about here]

Furthermore, the identity of each judge in charge of assessing the wines is known (first and last name), his or her gender, his or her industry credentials if any (no credential, Master Sommelier or Master of Wine) as well as his or her occupation in the industry: airline buyer, communicator (brand ambassador, journalist, etc.), consultant, educator, on-trade buyer (for bars, restaurants, etc., places selling beverages for immediate consumption on the premises), or off-trade buyer (for supermarkets, retailer, etc., when beverages are sold for consumption elsewhere, typically at home). The IWSC judges are presented as industry leaders and influencers, but some have closer ties to the wine business and wine consumers, such as on-trade buyers, than others, such as communicators or consultants.

The characteristics of the judges are shown in table 2.

[Insert table 2 about here]

The characteristics of the tasting teams are shown in table 3.

[Insert table 3 about here]

3.2 Differences among experts

Table 4 illustrates how individual ratings for the same wine can vary between judges within the same judging team. The score range, which measures the difference between the highest and lowest score, can be very wide, such as 9 points for wine #76750. The assessment takes place simultaneously, using identical glassware and in the same location, ensuring a consistent tasting context for all members of the judging team. The tasting environment cannot explain why the same wine receives different scores.

[Insert table 4 about here]

Table 5 presents the average individual score per gender, credential, and occupation, as well as a t-test for equality of means (given the size of the dataset and the law of large numbers, we assume a normal distribution of individual scores) on the first hand, and the average individual score according to the characteristics of the judging team on the other hand. On the whole, average scores vary significantly across judge's characteristics, except for master sommeliers and off-trade buyers. They also seem to vary according to the characteristics of the team.

[Insert table 5 about here]

4. Econometric analysis

4.1 Exploiting the panel structure of the dataset

The main objective of our analysis is to estimate a grade equation, at the judge level. As each wine is assessed simultaneously by several judges, our database has a typical panel structure. Although not longitudinal, we have a large number of units (the wines) for which we have two to seven observations (the members of the judging team), and the wine characteristics remain constant. Consequently, we incorporate wine fixed effects into the regression model, which enables us to account for unobserved heterogeneity and to mitigate potential sources of bias in the estimations. The right-hand side of our grade equation can focus only on the judge characteristics, without considering wine characteristics: gender (man – used as a reference – or woman), credential (no credential – used as a reference –, master sommelier, Master of Wine) and occupation (off-trade buyer – used as a reference –, on-trade buyer, airline buyer, communicator, consultant, educator) in Model 1. In terms of occupation, off-trade buyers are used as a reference point as they are closest to the consumer without mediation from a waiter or sommelier, unlike on-trade or airline buyers. Model 2 presents an alternative empirical specification where a dummy variable equals 1 if the judge is a buyer (airline, on-trade, or off-trade), and 0 otherwise.

The grade equation is estimated on the full dataset and on several subsets of wine styles and colours. The Huber-White estimator is used to obtain robust standard errors. The estimation results are presented in tables 6a (full dataset), 6b (sub-datasets by colour), and 6c (sub-datasets by wine style).

[Insert table 6a about here]

All F-test results converge and show the overall significance of our models and most of our results are significant at 1%.

Estimation results from both models are in accordance with each other (Table 6a). First, credentials do not matter to explain differences in wine scores. Neither Master sommeliers nor Master of Wine judges rate differently wines. Second, the regression highlights some differences across occupations. In Model 1, airline buyers, communicators, and educators give higher scores compared to off-trade buyers, while consultants and on-trade buyers are more severe. We can notice that the effect is particularly important for airline buyers, who give an average of 0.74 points more, all other things being equal. In model 2, we find that buyers, a variable that merges airline buyers, off-trade buyers, and on-trade buyers, give lower scores than judges from other occupations. Finally, it is worth noting that there is a gender bias in wine judging, with female judges consistently awarding lower scores than their male counterparts. This bias is statistically significant (at 1%) despite the relatively small effect size, with a penalty of -0.15 points in Model 1 and -0.115 points in Model 2.

The estimation results obtained on sub-datasets for red wines, white wines (Table 6b), and still wines (Table 6c) point in the same direction. However, sparkling wines appear to be an exception (Table 6c): Master of Wine judges are more severe than other judges, with a penalty of about -0.4 points. In comparison to off-trade buyers, judges from all other occupations are more generous, with almost one point more from airline buyers. These results confirm gender differences, with female judges having a stronger negative effect on sparkling wines (-0.250 points all else being equal).

[Insert table 6b about here]

[Insert table 6c about here]

4.2 Investigating peer effects

Another objective is to analyse peer effects among judges. For each individual score given by a judge, we know the composition of the team to which the judge belongs. Here, following Zimmerman (2003), who estimates the effect of roommate academic characteristics on individual grades, and Lavy and Schlosser (2011), who measure the effect of classroom gender ratios, we create a set of new variables to capture teammate characteristics: number of female judges among peers, number of Master of Wine judges and number of Master Sommelier judges, all in addition to the focal judge. We also include the size of each team (number of judges).

Taking into consideration the characteristics of each team, associated with a unique wine in the IWSC dataset, forces us to abandon wine fixed effects and to add wine characteristics in the grade equation. We include: the absence of vintage (1 if the wine has no vintage, 0 otherwise), the colour and style of the wine thanks to a series of dummy variables (orange, red, white, rosé on the one hand, still, fortified and sparkling on the other, with dry, medium and sweet as subcategories of still wines), whether the wine is made from a single variety or not (1 if it is a varietal wine, 0 otherwise), alcohol by volume, and country of origin (1 if the wine originates

from a given country, 0 otherwise, covering the OIV 2022 10 largest producing countries¹). Finally, as in the fixed effects models, the following characteristics of the focal judge are inserted into the empirical model: gender, credential and occupation.

The Huber-White estimator is used to obtain robust standard errors. Estimation results on the full dataset and on several subsets of wine styles and colours are presented in table 7.

[Insert table 7 about here]

Overall, the absence of vintage has a negative impact on the rating, with estimated coefficients ranging from -0.742 for white wines to -1.906 for still wines. Red wines receive a better rating than orange wines, and this is also the case for rosé and white wines in the sparkling subset. Fortified and sparkling wines are also rated better than still wines on average, and in the still wine subset, dry and medium wines are rated much lower than sweet wines. IWSC experts seem to prefer single-varietal wines, except for white wines (not significant) and sparkling wines (-0.278). Experts' ratings also increase with the alcohol content, all other things being equal, except for sparkling wines. In terms of country of origin, Chilean wines are rated lower than wines from the smallest wine-producing countries and, in the case of sparkling wines, all wines except those from France, the USA and Germany are penalised (the estimated coefficient is not significant for these three countries). Australian sparkling wines also benefit from a premium rating compared to sparkling wines from the smallest wine producing countries.

Interestingly, compared to the fixed effects models, the Master of Wine credential becomes significant in all specifications, with positive estimated coefficients ranging from 0.341 for red wines to 0.852 for still wines, and a negative estimated coefficient for sparkling wines (-0.478), while Master Sommelier remains non-significant. We also confirm the penalty applied by female judges, measuring a stronger gender effect for sparkling wines (-0.598). Occupation shows similar results as in the fixed effects models, although not systematically significant.

A number of team and peer effects have been identified. Firstly, when all other factors are held constant, the rating given by a judge tends to increase in line with the size of the team. Secondly, judges tend to be more severe when there are more female judges on the panel, with a marginal effect of every additional female judge in the team ranging from -0.110 for white wines to -0.483 for sparkling wines. The credentials of team members exert disparate effects upon the grades awarded to wines. While the addition of a Master of Wine judge to the panel results in an increase in the grades awarded to all wines, with the exception of sparkling wines, the inclusion of a Master Sommelier leads to a significant decrease in the grades awarded to sparkling wines only (-0.628).

5. Discussion and conclusion

The objective of this article is to elucidate the reasons behind the discrepancies in the scores assigned by experts to the same wines in identical tasting environments. Using a wide set of data, we focus on the personal characteristics of judges, including their gender, industry credentials, and occupation within the wine industry, and investigate how they respond to their peers' characteristics. Two sets of estimated grade equations, the first utilising wine fixed-

¹ A grade equation with dummy variables for the 40 countries in the dataset has also been estimated: it gives very similar results and increases the R² by 0.1 point. These estimation results are available from the authors on request.

effects models and the second examining peer effects, exhibit converging results across datasets, with the exception of sparkling wines.

In terms of occupation within the industry, compared to off-trade buyers, on-trade buyers and consultants tend to be more critical. Conversely, airline buyers, communicators, and educators tend to give higher grades than off-trade buyers. These results are consistent with recent literature indicating that the professional background and training of wine experts can explain differences in mental representation (Otheguy et al., 2021). This body of literature indicates that, depending on their background, experts can be oriented towards quality or style consistency (Joy et al., 2019), towards independent judgements targeting consumers (Grohmann et al., 2018), or towards a technical vs. appealing elaboration around their tasting experience (Honoré-Chazeau et al., 2024). In the course of personal discussions with some of the IWSC judges during the 2023 competition, we have gained the impression that off-trade buyers, more than other experts, adopt a business perspective when evaluating wines. This involves identifying those that will sell well. In essence, the knowledge that a medal often signifies a positive consumer response (Neuninger et al., 2017) may prompt a tendency towards greater generosity, given the extrinsic value of the award. Conversely, on-trade buyers understand the situation of being a waiter in front of consumers when they drink the wine and try not to be caught out with a wine. Consequently, they may be reluctant to be overly generous with a particular wine if they are uncertain about its quality.

Some credential effects also appear. The results of the wine fixed-effects models indicate that experts with no credentials and those with the titles of Master Sommelier or Master of Wine tend to provide similar scores for wines. Nevertheless, when examining the influence of team characteristics on judges, it becomes manifest that certain peer effects are linked to these credentials. The rating of the focal judge is found to increase with the number of Master of Wine judges on his team (with the exception of sparkling wines, where no significant effect is observed). Conversely, the rating of sparkling wines is found to decrease with the number of Master Sommelier judges. The Master of Wine certification is the most prestigious one in the wine industry (Orduña-Malea et al., 2021). Previous research has shown that reputation and status reinforce the cooperative behavior of other group members (Jazaieri et al., 2018) and status-based approaches of peer effects suggest that lower status individuals tend to adopt the behavior of higher status individuals (Cohen and Prinstein, 2006; Gerrard et al., 2008), what can explain the positive correlation measured here. Our results are more surprising regarding the negative effect of Master Sommelier judges belonging to the judging team, measured in the case of sparkling wines only. It reinforces the notion that sparkling wines represent a distinct category of wines. As noted by Thome and Paiva (2020, p. 37), “[s]parkling wine, which is known as a celebratory beverage, and also as a lifestyle symbol, has a strong symbolic function.” These symbolic aspects are probably also present in the expert's mind. Our findings are also consistent with Shepherd et al. (2023), who showed that experts conceptualise still wines and sparkling wines differently, as the later imply a specific core cerebral representation associated with terms such as champagne, bubbles and mousse.

Another important result is that female judges consistently give lower grades. The results of our study indicate that female judges are more likely to impose penalties than their male counterparts, what contradict the results obtained by Bodington and Malfeito-Ferreira (2018) with a smaller dataset. Furthermore, we show that judges, regardless of their own gender, tend to be more severe in their judgments when there are a greater number of female colleagues within their team. This phenomenon is very well documented in the court room (Boyd et al., 2010; Eren and Mocan, 2020). The critical mass hypothesis can explain this result, particularly given the progress women are making in the wine industry (Livat and Jaffré, 2022). Within a

group, individuals in the minority conform to the behaviour of the majority, which is viewed as the norm. As the minority reaches a critical mass, individuals feel less pressure and start making decisions that reflect their true opinion (Etzkowitz et al., 1994). In a similar vein, the propensity of women to make risky choices is intensified when they are surrounded by other women (Eren and Mocan, 2020). On the contrary, female judges may be hesitant to give high grades to wine due to a fear of failure. This could be a result of a lack of confidence in their own evaluations. This post hoc interpretation is consistent with Howel and Singer's (2019) findings on the self-confidence of male experts in a media context. Nevertheless, it cannot be ruled out that female judges are harsher because the wine deserves it, which raises some interesting new issues about how male judges rate wines.

The implications in terms of welfare are not immediately apparent. While it can be observed that female judges tend to be more severe than their male counterparts when assessing wines blindly, this does not imply that male judges are overly generous and do not discriminate wine quality. Furthermore, the International Wine and Spirits Competition, like other colleges of wine experts, pools the scores assigned by each panel of judges and publishes team grades, rather than individual grades. The penalty applied by female judges is, to some extent, mitigated by the premium given by male judges. Finally, given the feminisation of the industry, both in terms of producers and consumers, then the true expression of female experts' opinions could be beneficial in terms of social welfare.

One limitation of our analysis is the paucity of information about the cultural background of the experts, including their country of origin or nationality. It can be reasonably assumed that a wine judge coming from a wine-producing region may evaluate wines in a distinct manner compared to those originating from non-producing regions or regions with a nascent wine production history. Furthermore, the gender equality level of the country of origin is a pertinent issue to consider given the gender bias measured here. The age of the participants could also be a factor to take into account, as it may be positively correlated with their exposure to wine and the depth of their wine knowledge. This could suggest potential seniority effects, which are not investigated here. Furthermore, we have examined exogenous peer effects, yet have not considered the possibility of endogenous effects associated with the simultaneity of individual decisions. The rivalry effect (observing peers) and the image effect (being observed by peers) should be distinguished, but their elicitation presents substantial technical challenges (Beugnot et al., 2013). Recently, researchers have also focused on tasting narratives and tasting notes to investigate gender issues in wine expertise (Honoré-Chedozeau et al., 2024; Masset et al., 2024). Further research can delve deeper into textual and sentiment analysis, using artificial intelligence to classify wine descriptors as feminine or masculine, assess their valence, and examine their link with expert gender.

Acknowledgement

We are grateful to the editor and to two anonymous reviewers for their insightful comments and suggestions. We would like to express our gratitude to Christelle Guibert for granting us access to the IWSC 2022 dataset and to Jeremy Cukierman MW for his valuable comments. The usual disclaimers apply.

REFERENCES

- Anderson, M., & Magruder, J. (2012). Learning from the crowd: Regression discontinuity estimates of the effects of an online review database. *The Economic Journal*, 122(563), 957-989.
- Ashton, R. H. (2012). Reliability and consensus of experienced wine judges: Expertise within and between?. *Journal of Wine Economics*, 7(1), 70-87.
- Ashton, R. H. (2017). Dimensions of expertise in wine evaluation. *Journal of Wine Economics*, 12(1), 59-83.
- Beckmann, D., & Menkhoff, L. (2008). Will women be women? Analyzing the gender difference among financial experts. *Kyklos*, 61(3), 364-384.
- Berg, E. C., Mascha, M., & Capehart, K. W. (2022). Judging reliability at wine and water competitions. *Journal of Wine Economics*, 17(4), 311-328.
- Beugnot, J., Fortin, B., Lacroix, G., & Villeval, M. C. (2013). Social networks and peer effects at work. *CIRANO-Scientific Publications 2013s-27*.
- Bodington, J., & Malfeito-Ferreira, M. (2018). Do female and male judges assign the same ratings to the same wines? Large sample results. *Journal of wine economics*, 13(4), 403-408.
- Bonnet, C., Hilger, J., & Villas-Boas, S. B. (2020). Reduced form evidence on belief updating under asymmetric information—consumers’ response to wine expert opinions. *European Review of Agricultural Economics*, 47(5), 1668-1696.
- Boyd, C. L., Epstein, L., & Martin, A. D. (2010). Untangling the causal effects of sex on judging. *American journal of political science*, 54(2), 389-411.
- Brochet, F., & Dubourdieu, D. (2001). Wine descriptive language supports cognitive specificity of chemical senses. *Brain and language*, 77(2), 187-196.
- Cardebat, J. M., & Livat, F. (2016). Wine experts’ rating: a matter of taste?. *International Journal of Wine Business Research*, 28(1), 43-58.
- Cawley, D. (2018). The Power of Wine Language - Critics, Labels and Sexism. *Dublin Gastronomy Symposium*, Dublin, Ireland, 29-31 May. doi.org/10.21427/g5a7-x024.
- Chocarro, R., & Cortinas, M. (2013). The impact of expert opinion in consumer perception of wines. *International Journal of Wine Business Research*, 25(3), 227-248.

- Chu, S., and Downes, J. J., (2000), "Long live Proust: The odour-cued autobiographical memory bump", *Cognition*, 75(2), B41-B50.
- Cliff, M. A., & King, M. C. (1996). A proposed approach for evaluating expert wine judge performance using descriptive statistics. *Journal of Wine Research*, 7(2), 83-90.
- Cohen, G. L., & Prinstein, M. J. (2006). Peer contagion of aggression and health risk behavior among adolescent males: An experimental investigation of effects on public conduct and private attitudes. *Child development*, 77(4), 967-983.
- Combris, P., Lecocq, S., & Visser, M. (1997). Estimation of a hedonic price equation for Bordeaux wine: does quality matter?. *The Economic Journal*, 107(441), 390-402.
- D'Alessandro, S., & Pecotich, A. (2013). Evaluation of wine by expert and novice consumers in the presence of variations in quality, brand and country of origin cues. *Food Quality and Preference*, 28(1), 287-303.
- Depetris Chauvin, N., Valentin, D., Behrens, J. H., & Rodrigues, H. (2024). Country-of-Origin as bias inducer in experts' wine judgments-A sensory experiment in a world wine fair. *International Journal of Gastronomy and Food Science*, 100883.
- Eren, O., & Mocan, N. H. (2020). *Judge peer effects in the courthouse* (No. w27713). National Bureau of Economic Research.
- Etzkowitz, H., Kemelgor, C., Neuschatz, M., Uzzi, B., & Alonzo, J. (1994). The paradox of critical mass for women in science. *Science*, 266(5182), 51-54.
- Falk, A., & Hermle, J. (2018). Relationship of gender differences in preferences to economic development and gender equality. *Science*, 362(6412), eaas9899.
- Fischer, U., & Loewe-Stanienda, B. (2000). Impact of wine glasses for sensory evaluation. *Journal International des Sciences de la Vigne et du Vin*, 33, 71-80.
- Gawel, R., & Godden, P. W. (2008). Evaluation of the consistency of wine quality assessments from expert wine tasters. *Australian journal of grape and wine research*, 14(1), 1-8.
- Gerrard, M., Gibbons, F. X., Houlihan, A. E., Stock, M. L., & Pomery, E. A. (2008). A dual-process approach to health risk decision making: The prototype willingness model. *Developmental review*, 28(1), 29-61.
- Grohmann, B., Peña, C., & Joy, A. (2018). Wine quality and sensory assessments: do distinct local groups of wine experts differ?. *Journal of wine research*, 29(4), 278-289.
- Hendricks, J. L., Call, M. L., & Campbell, E. M. (2023). High performer peer effects: A review, synthesis, and agenda for future research. *Journal of Management*, 49(6), 1997-2029.
- Hilger, J., Rafert, G., & Villas-Boas, S. (2011). Expert opinion and the demand for experience goods: an experimental approach in the retail wine market. *Review of Economics and Statistics*, 93(4), 1289-1296.
- Holden, R., Keane, M., & Lilley, M. (2021). Peer effects on the united states supreme court. *Quantitative Economics*, 12(3), 981-1019.

- Honoré-Chedozeau, C., Otheguy, M., & Valentin, D. (2024). Tell us how you taste wine, and we will tell you what kind of expert you are!. *Food Research International*, 178, 113899.
- Howell, L., & Singer, J. B. (2019). Pushy or a princess? Women experts and UK broadcast news. *Journalism Practice*, 13(8), 1018-1023.
- Humphreys, A., & Carpenter, G. S. (2018). Status games: Market driving through social influence in the US wine industry. *Journal of Marketing*, 82(5), 141-159.
- Jaud, D. A., Gergaud, O., & Lunardo, R. (2023). Family and peer communication and wine consumption among young adults: Examining the role of responsible drinking practices. *British Food Journal*, 125(6), 2070-2086.
- Jazaieri, H., Logli Allison, M., Campos, B., Young, R. C., & Keltner, D. (2019). Content, structure, and dynamics of personal reputation: The role of trust and status potential within social networks. *Group Processes & Intergroup Relations*, 22(7), 964-983.
- Joy, A., Grohmann, B., & Pena, C. (2019). Preliminary thoughts on the importance of sensory profiling for strategic decision-making in Canadian wineries. In *Case Studies in the Wine Industry* (pp. 167-179). Woodhead Publishing.
- Kaimann, D., & Bru, C. L. M. S. (2024). Sounds too feminine? Blind tastings, phonetic gender scores, and the impact on professional critics. *Journal of Wine Economics*, 1-20.
- Komal, B., Ezeani, E., Usman, M., Kwabi, F., & Ye, C. (2023). Do the educational profile, gender, and professional experience of audit committee financial experts improve financial reporting quality?. *Journal of International Accounting, Auditing and Taxation*, 53, 100580.
- Lavy, V., & Schlosser, A. (2011). Mechanisms and impacts of gender peer effects at school. *American Economic Journal: Applied Economics*, 3(2), 1-33.
- Lesschaeve, I. (2007). Sensory evaluation of wine and commercial realities: Review of current practices and perspectives. *American Journal of Enology and Viticulture*, 58(2), 252-258.
- Livat, F., & Jaffré, C. (2022). Women in wine... occasionally: Gendered roles in the wine industry. *The Routledge Handbook of Wine and Culture*, 320-330.
- Livat, F., Marks, D., & Remaud, H. (2024). Wining and dining out: on understanding restaurant wine lists. *Food, Culture & Society*, 1-20.
- Livat, F., & Remaud, H. (2021). Do wine judges give higher scores to wines made with less-known grape varieties?. *International Journal of Entrepreneurship and Small Business*, 44(2), 106-117.
- Lizzeri, A. (1999). Information revelation and certification intermediaries. *The RAND Journal of Economics*, 214-231.
- Lombardo, M., Amoah, I., Feraco, A., Guseva, E., Gorini, S., Camajani, E., ... & Caprio, M. (2023, October). Gender Differences in Taste and Food Habits: Implications for Personalized Nutrition. In *Biology and Life Sciences Forum*, 29 (1), 1.

- Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The review of economic studies*, 60(3), 531-542.
- Marlowe, B., Brown, E. A., Schrier, T., & Zheng, T. (2017). Beverage bloggers: A developing relationship between wine blogger expertise and twitter followers. *International Journal of Hospitality Beverage Management*, 1(1), 2.
- Martín-Román, Á., Moral, A., & Martínez-Matute, M. (2015). Peer effects in judicial decisions: Evidence from Spanish labour courts. *International Review of Law and Economics*, 42, 20-37.
- Masset, P., Terrier, L., & Livat, F. (2023). Can a wine be feminine? Gendered wine descriptors and quality, price, and aging potential. *Journal of Wine Economics*, 1-13.
- Miles, T. J. (2012). The law's delay: A test of the mechanisms of judicial peer effects. *Journal of Legal Analysis*, 4(2), 301-327.
- Mora, M., Urdaneta, E., & Chaya, C. (2018). Emotional response to wine: Sensory properties, age and gender as drivers of consumers' preferences. *Food Quality and Preference*, 66, 19-28.
- Neuninger, R., Mather, D., & Duncan, T. (2017). Consumer's scepticism of wine awards: A study of consumers' use of wine awards. *Journal of Retailing and Consumer Services*, 35, 98-105.
- Nunes, P., Muxagata, S., Correia, A. C., Nunes, F. M., Cosme, F., & Jordão, A. M. (2017). Effect of oak wood barrel capacity and utilization time on phenolic and sensorial profile evolution of an Encruzado white wine. *Journal of the Science of Food and Agriculture*, 97(14), 4847-4856.
- OIV (2022). State of the world vine and wine sector in 2022. Available online: www.oiv.int/sites/default/files/documents/OIV_State_of_the_world_Vine_and_Wine_sector_in_2022_2.pdf [accessed on January 29, 2024].
- Orduña Malea, E., Font-Julian, C. I., Ontalba Ruipérez, J. A., & Compes López, R. (2021). Masters of Wine on Twitter: presence, activity, impact and community structure. *Wine Economics and Policy*, 10(1), 73-88.
- Otheguy, M., Honoré-Chedozeau, C., & Valentin, D. (2021). Do wine experts share the same mental representation? A drawing elicitation study with wine makers, sellers, and critics. *Food Quality and Preference*, 94, 104302.
- Paroissien, E., & Visser, M. (2020). The Causal Impact of Medals on Wine Producers' Prices and the Gains from Participating in Contests. *American Journal of Agricultural Economics*, 102(4), 1135-1153.
- Parr, W. V., Green, J. A., & White, K. G. (2006). Wine judging, context and New Zealand Sauvignon blanc. *European Review of Applied Psychology*, 56(4), 231-238.
- Parr, W. V., Green, J. A., White, K. G., & Sherlock, R. R. (2007). The distinctive flavour of New Zealand Sauvignon blanc: Sensory characterisation by wine professionals. *Food quality and preference*, 18(6), 849-861.

- Parr, W.V., (2019), “Demystifying wine tasting: Cognitive psychology's contribution”, *Food Research International*, 124, 230-233.
- Pérez-Jiménez, M., Muñoz-González, C., Chaya, C., Fernández-Ruiz, V., Álvarez, M. D., Herranz, B., & Pozo-Bayon, M. A. (2022). Insights on the effect of age and gender on in-mouth volatile release during wine tasting. *Food Research International*, 155, 111100.
- Portyanko, S., Reinmoeller, P., Hussels, S., & Turner, N. (2023). Peer effects and intentional entrepreneurial behaviour: A systematic literature review and research agenda. *International Journal of Management Reviews*, 25(3), 515-545.
- Rodrigues, H., & Parr, W. V. (2019). Contribution of cross-cultural studies to understanding wine appreciation: A review. *Food research international*, 115, 251-258.
- Rodrigues, H., Richards, A., & Carvalho, F. M. (2023). Wine tasters' sensory response to delicate and robust flower arrangements. *International Journal of Gastronomy and Food Science*, 34, 100783.
- Sáenz-Navajas, M.P., Avizcuri, J.M., Echávarri, J.F., Ferreira, V., Fernández-Zurbano, P., Valentin, D., (2016), “Understanding quality judgements of red wines by experts: Effect of evaluation condition”, *Food Quality and Preference*, 48, 216-227.
- Scaman, C. H., Dou, J., Cliff, M. A., Yuksel, D., & King, M. C. (2001). Evaluation of wine competition judge performance using principal component similarity analysis. *Journal of Sensory Studies*, 16(3), 287-300.
- Sena-Esteves, M. M., Mota, M., & Malfeito-Ferreira, M. (2018). Patterns of sweetness preference in red wine according to consumer characterisation. *Food Research International*, 106, 38-44.
- Shepherd, H., Parr, W. V., Monaco, G. L., & Rodrigues, H. (2023). The meaning of the word elegance as a wine descriptor: Effect of expertise and wine type. *Food Research International*, 164, 112399.
- Songer, D. R., & Crews-Meyer, K. A. (2000). Does judge gender matter? Decision making in state supreme courts. *Social Science Quarterly*, 750-762.
- Spence, C. (2020). Wine psychology: basic & applied. *Cognitive research: principles and implications*, 5(1), 1-18.
- Spence, C., Velasco, C., & Knoeferle, K. (2014). A large sample study on the influence of the multisensory environment on the wine drinking experience. *Flavour*, 3(1), 1-12.
- Suárez, A., Rodrigues, H., Williams, S., Fernández-Zurbano, P., Depetris-Chauvin, N., Dunn, G., Sáenz-Navajas, M.P., (2023), “Effect of culture and familiarity on wine perception: a study with British and Spanish wine experts”, *OENO One*.
- Valentin, D., Parr, W. V., Peyron, D., Grose, C., & Ballester, J. (2016). Colour as a driver of Pinot noir wine quality judgments: An investigation involving French and New Zealand wine professionals. *Food Quality and Preference*, 48, 251-261.

Villeval, M. C. (2020). Performance feedback and peer effects: A review. *Available at SSRN* 3543371.

Weerasekara, N., & Streletskaia, N. A. (2024). Tasting and consumer demand for wine: do peers and experts matter?. *Agricultural and Resource Economics Review*, 53(1), 66-86.

Wilkinson, I. A., & Fung, I. Y. (2002). Small-group composition and peer effects. *International journal of educational research*, 37(5), 425-447.

Zalata, A. M., Tauringana, V., & Tingbani, I. (2018). Audit committee financial expertise, gender, and earnings management: Does gender of the financial expert matter?. *International review of financial analysis*, 55, 170-183.

Zimmerman, D. J. (2003). Peer effects in academic outcomes: Evidence from a natural experiment. *Review of Economics and statistics*, 85(1), 9-23.

Table 1: Characteristics of the wines – IWSC 2022

Wine characteristics	Obs	Mean or %	Std. Dev.	Min	Max
Percent alcohol	18,224	13.405	1.506	5	28
Single varietal	10,631	58.34			
No vintage	1,872	10.27			

Style:

Still	15,464	84.86
-------	--------	-------

Sparkling	2,184	11.98
-----------	-------	-------

Fortified	576	3.16
-----------	-----	------

Color:

Red	9,728	53.38
-----	-------	-------

White	6,926	38.00
-------	-------	-------

Rosé	1,49	8.18
------	------	------

Orange	80	0.44
--------	----	------

Sweetness:

Dry	13,723	75.30
-----	--------	-------

Medium	1,455	7.98
--------	-------	------

Sweet	286	1.57
-------	-----	------

*Country of origin**(OIV 2022 10 largest producing countries):*

Italy	3,315	18,19
-------	-------	-------

France	2,537	13,92
--------	-------	-------

Spain	2,430	13,33
-------	-------	-------

USA	532	2,92
Australia	1706	9,36
Chile	526	2,89
Argentina	818	4,49
South Africa	368	2,02
Germany	72	0,40
Portugal	1229	6,74

Table 2: Characteristics of the judges – IWSC 2022

	Mean	Std. Dev.	%	Min	Max
Number of wines rated	195.957	142.289		43	605
Woman			33.9	0	1
Master Sommelier			12.2	0	1
Master of Wine			10.5	0	1
Airline buyer			0.9	0	1
On-trade buyer			34.9	0	1
Off-trade buyer			48.1	0	1
Communicator			7.0	0	1
Educator			8.7	0	1
Consultant			0.4	0	1

Table 3: Characteristics of the judging teams – IWSC 2022

	Mean per team	% of teams	Std. Dev.	Min	Max
Number of judges	3.458		0.558	2	7
2 judges		0.04			
3 judges		56.13			

4 judges		42.80		
5 judges		0.03		
6 judges		0.95		
7 judges		0.04		
Number of female judges	1.172		0.758	0 3
No female judge		19.34		
1 female judge		46.37		
2 female judges		32.02		
3 female judges		2.26		
Number of MW* judges	0.355		0.517	0 2
No MW judge		66.37		
1 MW judge		31.72		
2 MW judges		1.91		
Number of MS* judges	0.416		0.607	0 3
No MS judge		63.36		
1 MS judge		32.88		
2 MS judges		2.51		

3 MS judges	1.25
-------------	------

* MW: Master of wine / MS: Master Sommelier

Table 4: Sample of individual ratings over 100 points

(source: IWSC 2022)

Wine ID	Judge A	Judge B	Judge C	Judge D	Score range
69732	85	92	85	84	8
76750	84	86	82	91	9
82237	87	83	86	86	4

Table 5: Average score, judge, and team characteristics

Judge characteristics	Average score over 100 points	T-test equality of means*	Pr.
<i>Gender:</i>			
Male	87.32	7.7509	0.000
Female	86.94		
<i>Credential:</i>			
No credential	87.16	-2.7798	0.005
Master sommelier	87.15	0.7119	0.476
Master of Wine	87.51	-4.5626	0.000
<i>Occupation:</i>			
Airline buyer	87.61	-1.7425	0.081
On-trade buyer	87.02	5.4992	0.000
Off-trade buyer	87.22	-1.0607	0.289
Communicator	87.41	-2.6233	0.008
Educator	87.58	-5.1112	0.000

Consultant	85.97	3.1324	0.00	2
Team characteristics	Average score over 100 points	Std. Dev.	Min	Max
<i>Number of judges:</i>				
2 judges	83	6.391	70	90
3 judges	87.178	3.248	40	99
4 judges	87.216	3.049	70	98
5 judges	87.8	1.643	85	89
6 judges	87.218	2.023	83	92
7 judges	88.286	0.488	88	89
<i>Number of female judges:</i>				
No female judge	87.549	3.455	50	99
1 female judge	87.264	3.095	40	98
2 female judges	86.941	3.036	70	98
3 female judges	86.262	2.945	80	96
<i>Number of MW** judges:</i>				
No MW judge	86.992	3.094	40	98
1 MW judge	87.636	3.282	50	99

2 MW judges	86.845	2.332	80	92
<i>Number of MS** judges:</i>		0.607	0	3
No MS judge	87.225	3.152	50	99
1 MS judge	87.154	3.189	40	98
2 MS judges	87.202	2.974	75	97
3 MS judges	86.588	2.79	79	93

* Test for the focal characteristics vs. other characteristics within the characteristic category as a whole.

** MW: Master of wine / MS: Master Sommelier.

Table 6a: Grade equation estimation results – Full dataset

	Model 1		Model 2	
	Estimated coef.	Std. error	Estimated coef.	Std. error
No credential	ref.		ref.	
Master sommelier	0.0181	(0.0477)	-0.0239	(0.0457)
Master of Wine	-0.0514	(0.0575)	-0.0825	(0.0512)
Off-trade buyer	ref.		ref.	
Airline buyer	0.740***	(0.201)		
Communicator	0.289***	(0.0596)		
Consultant	-0.411***	(0.135)		
Educator	0.176***	(0.0590)		
On-trade buyer	-0.0875**	(0.0358)		
Buyer			-0.243***	(0.0419)
Gender (woman=1)	-0.150***	(0.0345)	-0.115***	(0.0336)
Constant	87.24***	(0.0220)	87.45***	(0.0373)
Wine fixed effect	Yes		Yes	
Observations	18,218		18,218	
F (prob.)	11.54 (0.000)		13.02 (0.000)	

Number of wines	5,395	5,395
-----------------	-------	-------

Robust standard errors in parentheses

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Journal Pre-proofs

Table 6b: Grade equation estimation results – Sub-datasets by color

	Red wines				White wines			
	<i>Model 1</i>		<i>Model 2</i>		<i>Model 1</i>		<i>Model 2</i>	
	Est. coef.	Std. err.	Est. coef.	Std. err.	Est. coef.	Std. err.	Est. coef.	Std. err.
No credential	ref.		ref.		ref.		ref.	
Master sommelier	-0.00900	(0.0672)	-0.0367	(0.0643)	0.0368	(0.0760)	-0.0174	(0.0726)
Master of Wine	-0.0725	(0.0876)	-0.112	(0.0769)	-0.0557	(0.0858)	-0.0689	(0.0777)
Off-trade buyer	ref.		ref.		ref.		ref.	
Airline buyer	0.679**	(0.306)			0.729**	(0.301)		
Communicator	0.318***	(0.0849)			0.246***	(0.0954)		
Consultant	-0.503***	(0.186)			-0.369	(0.253)		
Educator	0.171**	(0.0869)			0.206**	(0.0946)		

On-trade buyer	-0.0477 (0.0498)		-0.122** (0.0618)	
Buyer		-0.238*** (0.0606)		-0.252*** (0.0669)
Gender (woman=1)	-0.113** (0.0483)	-0.0835* (0.0470)	-0.170*** (0.0575)	-0.136** (0.0564)
Constant	87.28*** (0.0309)	87.50*** (0.0534)	87.24*** (0.0365)	87.45*** (0.0597)
Wine fixed effect	Yes	Yes	Yes	Yes
Observations	9,726	9,726	6,922	6,922
F (prob.)	5.36 (0.000)	5.68 (0.000)	4.57 (0.000)	5.47 (0.000)
Number of wines	3,062	3,062	2,279	2,279

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 6c: Grade equation estimation results – Sub-datasets by wine style

	Still wine				Sparkling wine			
	<i>Model 1</i>		<i>Model 2</i>		<i>Model 1</i>		<i>Model 2</i>	
	Est. coef.	Std. err.	Est. coef.	Std. err.	Est. coef.	Std. err.	Est. coef.	Std. err.
No credential	ref.		ref.		ref.		ref.	
Master sommelier	-0.00484	(0.0510)	-0.0582	(0.0486)	-0.0745	(0.144)	-0.0439	(0.145)
Master of Wine	0.0898	(0.0654)	0.0723	(0.0601)	-0.372***	(0.130)	-0.447***	(0.110)
Off-trade buyer	ref.		ref.		ref.		ref.	
Airline buyer	0.756**	(0.314)			0.934***	(0.297)		
Communicator	0.271***	(0.0604)			0.515***	(0.169)		

Consultant	-0.402*** (0.136)		No obs.	No obs.
Educator	0.0926 (0.0652)		0.481*** (0.161)	
On-trade buyer	-0.0975** (0.0383)		0.210* (0.112)	
Buyer		-0.194*** (0.0437)		-0.454*** (0.122)
Gender (woman=1)	-0.168*** (0.0384)	-0.150*** (0.0379)	-0.250*** (0.0933)	-0.177** (0.0896)
Constant	87.17*** (0.0236)	87.33*** (0.0403)	87.04*** (0.0724)	87.56*** (0.0923)
Wine fixed effect	Yes	Yes	Yes	Yes
Observations	15,458	15,458	2,184	2,184
F (prob.)	9.30 (0.000)	9.83 (0.000)	7.30 (0.000)	8.23 (0.000)
Number of wines	4,517	4,517	691	691

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 7: Grade equation with peer effects: Estimation results

	All wines		Red wines		White wines		Still wines		Sparkling wines	
	Est. coef.	Std. Err.	Est. coef.	Std. Err.	Est. coef.	Std. Err.	Est. coef.	Std. Err.	Est. coef.	Std. Err.
<i>Wine characteristics:</i>										
No vintage	- 1.044***	(0.116)	- 1.162***	(0.167)	- 0.742***	(0.207)	- 1.906***	(0.205)	- 0.808***	(0.138)
Orange wine	Ref.						Ref.		Ref.	
Red wine	0.566*	(0.332)					0.629*	(0.322)	1.229	(0.762)

Rosé wine	0.250	(0.338)			0.254	(0.330)	1.906**	(0.780)
White wine	0.487	(0.332)			0.521	(0.323)	1.393*	(0.760)
Still wine	Ref.		Ref.		Ref.			
Fortified wine	2.258***	(0.243)	2.524***	(0.326)	1.840***	(0.420)		
Sparkling wine	0.537***	(0.111)	0.570***	(0.162)	0.228	(0.186)		
Sweet wine (still only)							Ref.	
Dry wine (still only)							-	2.658*** (0.230)
Medium wine (still only)							-	3.091*** (0.239)
Single varietal	0.240***	(0.0487)	0.287***	(0.0660)	0.0782	(0.0828)	0.342***	(0.0518)
Alcohol by volume	0.243***	(0.0242)	0.220***	(0.0319)	0.292***	(0.0414)	0.320***	(0.0265)
							-0.0193	(0.0906)

Italy	0.0874	(0.0726)	-0.0278	(0.101)	0.321***	(0.124)	0.572***	(0.0775)	-	1.937***	(0.246)
France	1.300***	(0.0800)	1.183***	(0.115)	1.485***	(0.127)	1.591***	(0.0889)	0.00958	(0.192)	
Spain	0.0418	(0.0780)	0.0306	(0.102)	0.0352	(0.138)	0.280***	(0.0852)	-	1.018***	(0.291)
USA	0.599***	(0.169)	0.559***	(0.217)	0.755***	(0.281)	0.614***	(0.171)	-0.442	(0.977)	
Australia	-	(0.0826)	0.238***	(0.111)	-0.150	(0.140)	-0.311**	(0.0834)	-0.215**	1.331***	(0.500)
Chile	-	(0.129)	1.471***	(0.182)	-	(0.224)	1.384***	(0.130)	-	4.303***	(0.651)
Argentina	0.559***	(0.104)	0.603***	(0.139)	0.572***	(0.168)	0.746***	(0.107)	-	3.497***	(0.356)
South Africa	0.380*	(0.195)	0.875***	(0.268)	0.0862	(0.323)	0.865***	(0.209)	-	2.133***	(0.462)
Germany	1.965***	(0.396)	1.056*	(0.601)	2.215***	(0.560)	2.324***	(0.479)	0.277	(0.686)	

Portugal	0.307***	(0.102)	0.368***	(0.143)	0.189	(0.164)	0.887***	(0.106)	-	1.634***	(0.405)
Other country of origin	Ref.		Ref.		Ref.		Ref.		Ref.		
<i>Focal judge characteristics:</i>											
No credentials	Ref.		Ref.		Ref.		Ref.		Ref.		
Master of Wine	0.376***	(0.0987)	0.341**	(0.143)	0.439***	(0.155)	0.852***	(0.112)	-0.478*	(0.246)	
Master sommelier	-0.0764	(0.0746)	-0.104	(0.104)	-0.0422	(0.118)	-0.00648	(0.0787)	-0.687**	(0.318)	
Gender (woman=1)	-0.407***	(0.0529)	-0.526***	(0.0714)	-0.248***	(0.0893)	-0.430***	(0.0561)	-0.598***	(0.180)	
Off-trade buyer	Ref.		Ref.		Ref.		Ref.		Ref.		
Airline buyer	0.377	(0.282)	0.401	(0.405)	0.219	(0.441)	0.547	(0.422)	0.814**	(0.394)	
Communicator	0.172*	(0.103)	0.188	(0.147)	0.0922	(0.161)	0.199*	(0.114)	0.349	(0.261)	

Consultant	-	0.791***	(0.293)	-0.536	(0.415)	1.369***	(0.431)	-0.649**	(0.294)	No obs.
Educator		0.241**	(0.100)	0.142	(0.142)	0.383**	(0.161)	-0.00374	(0.112)	0.783*** (0.228)
Ontrade buyer		-0.0210	(0.0538)	-0.0591	(0.0734)	0.0403	(0.0894)	-0.0584	(0.0566)	0.101 (0.179)
<i>Team characteristics:</i>										
Number of team members		0.266***	(0.0449)	0.369***	(0.0601)	0.137*	(0.0806)	0.156***	(0.0472)	0.128 (0.271)
Number of female judges	-	0.291***	(0.0375)	0.438***	(0.0505)	-0.110*	(0.0633)	0.252***	(0.0397)	0.483*** (0.144)
Number of Master of Wine judges		0.493***	(0.0552)	0.465***	(0.0761)	0.577***	(0.0951)	0.738***	(0.0629)	0.0306 (0.173)
Number of Master Sommelier judges		-0.0588	(0.0467)	-0.0697	(0.0638)	-0.0160	(0.0750)	0.0368	(0.0486)	-0.628** (0.257)
Constant		82.34***	(0.479)	83.05***	(0.455)	82.40***	(0.621)	83.91***	(0.514)	87.44*** (1.582)

Observations	18,218	9,726	6,922	15,458	2,184
R-squared	0.080	0.086	0.076	0.083	0.112

Robust standard errors in parentheses.

*** p<0.01, ** p<0.05, * p<0.1

Journal Pre-proofs

**Factors explaining differences in wine experts' ratings:
the case of gender, credentials, occupation and peer effects**

Florine Livat, Hervé Remaud and Marta Fernández-Olmos

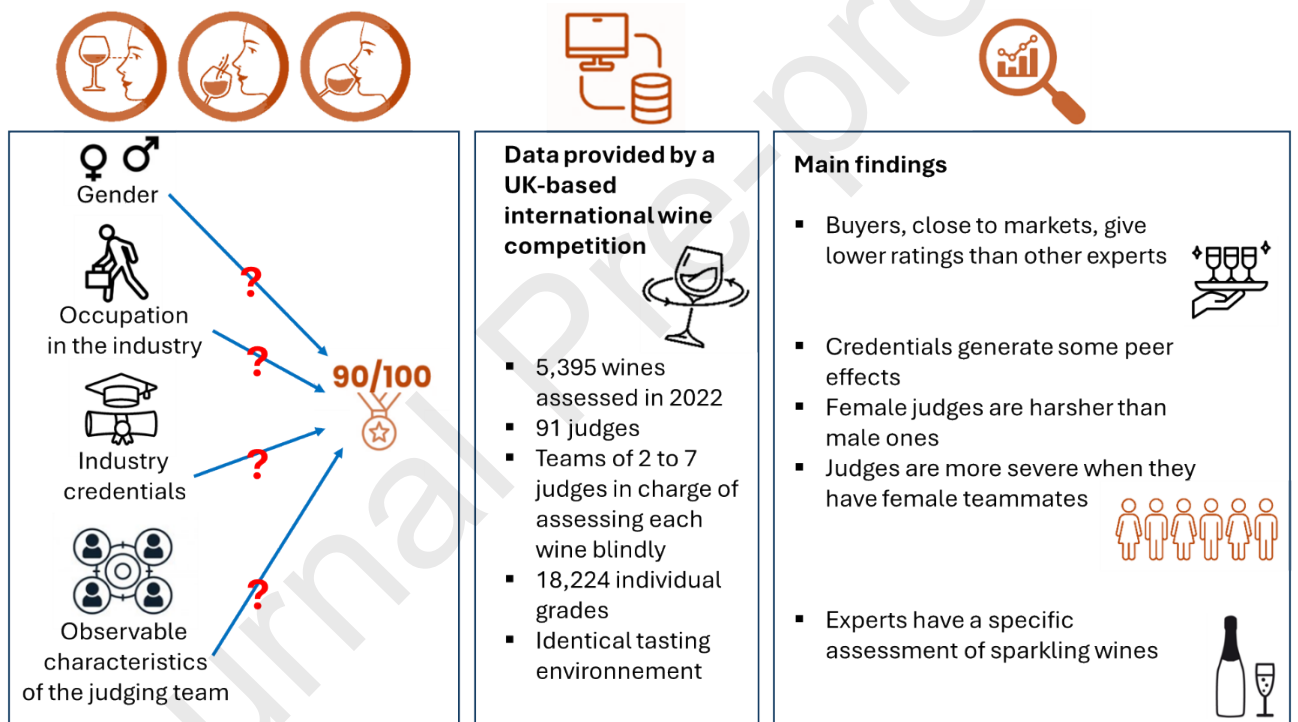
Highlights

- Buyers, close to markets, give lower ratings than other experts.
- Highly reputed industry credentials generate some peer effects.
- Female judges consistently give lower grades than their male counterparts.
- Judges tend to be harsher when there are more female peers, regardless of their own gender.
- Experts have a specific assessment of sparkling wines.

**Factors explaining differences in wine experts' ratings:
the case of gender, credentials, occupation and peer effects**

Florine Livat, Hervé Remaud and Marta Fernández-Olmos

Graphical abstract



June 17, 2024

Manuscript ID: **FOODRES-D-24-01180**

Dear Food Research International editor in chief,

Dear Special Issue editors,

We declare that the manuscript entitled “Factors explaining differences in wine experts’ ratings: the case of gender, credentials, occupation and peer effects” written by Hervé Remaud, Marta Fernández-Olmos and myself, is original, has not been published before and is not currently being considered for publication elsewhere. Furthermore, we did not utilize generative AI to compose this article, except to enhance its language and readability (DeepL Write). After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

We know of no conflicts of interest associated with this publication, and there has been no significant financial support for this work that could have influenced its outcome. As Corresponding Author, I confirm that the manuscript has been read and approved for submission by all the named authors.

We hope that our revision will be found to be both convincing and relevant for the FRI special issue. We look forward to hearing from you.

Sincerely,

Florine Livat, PhD

A handwritten signature in black ink, appearing to be 'FLORINE LIVAT', written in a cursive style.

KEDGE Business School, 680 Cours de la Libération, 33400 Talence, France

florine.livat@kedgebs.com / +33 684 113 661