CARMA 2026

8th Int. Conf. on Advanced Research Methods and Analytics
1-3 July 2026, Valencia, Spain
DOI: https://doi.org/10.4995/CARMA2026.2026.22029

Uncorking Profits: A Causal Analysis of Enotourism on Firm Performance

Rita Jiménez[1], Ana Debón[1], Josep Domenech[1]
[1]eSMART Research Center, Universitat Politècnica de València, Spain.

How to cite: Jiménez, R.; Debón, A.; Domenech, J. (2026). Uncorking Profits: A Causal Analysis of Enotourism on Firm Performance. In: 8th International Conference on Advanced Research Methods and Analytics (CARMA 2026). Valencia, 1-3 July 2026. https://doi.org/10.4995/CARMA2026.2026.22029

Abstract

This paper investigates the causal impact of enotourism on the financial performance of wine-producing firms, addressing the endogeneity of this strategic decision. Unlike previous empirical literature, this study applies an Instrumental Variables approach using a Two-Stage Least Squares model on a sample of 688 Spanish wineries in 2024. The results demonstrate that offering enotourism services has a direct, positive, and statistically significant effect on a winery's value added. Furthermore, firm size emerges as a critical internal determinant of profitability. This research provides robust quantitative evidence establishing enotourism as a genuine driver of economic performance, highlighting the critical need to control for self-selection bias in wine economics.

Keywords: *Wine; Enotourism; Instrumental Variables: Causality; Endogenity.*

1. Introduction

Wine tourism has established itself as a strategic activity of growing popularity, prompting numerous studies on its economic and territorial impact (Gómez et al., 2019). However, much of the previous empirical literature analyses this activity by assuming exogeneity, treating the provision of these services as a randomly assigned event independent of the firm's unobservable characteristics. This approach ignores the fact that the decision to offer enotourism is a strategic, endogenous choice made by the firm. Since it is not a random control variable, but rather a conscious business decision conditioned by internal factors, traditional estimation suffers from a self-selection bias that must be addressed with methodological rigor.

This paper investigates the causal impact of enotourism on the financial performance of wineries. Given that the adoption of this activity is an internal decision prone to endogeneity, estimation via an instrumental variables (IV) model is proposed. This econometric approach is ideal for correcting selection bias, controlling for unobserved heterogeneity, and consistently isolating the causal effect of enotourism on corporate profits (Angrist & Pischke, 2008;

Article in Press

This work is licensed under a Creative Commons License CC BY-NC-SA 4.0
Editorial Universitat Politècnica de València

Wooldridge, 2010). Although the sample is limited to Spain, the implications may be relevant for both Old World and New World wine regions.

2. Related work

The empirical literature has traditionally examined enotourism as an experiential marketing instrument oriented toward competitive differentiation. Seminal contributions by Bruwer et al. (2013) and Frost et al. (2020) demonstrate that that winery experiences strengthen brand attachment and encourage repeat visits. More broadly, enotourism may also enhance reputation and demand traction at the firm level (Winfree et al., 2018).

More recently, the analytical focus has shifted toward the sector's digital ecosystem. Emerging evidence suggests that user-generated evaluations (eWOM) exert a greater influence on wine price formation than traditional expert assessments (Bazen et al., 2024). Despite the salience of this dynamic, the quantitative modelling of the relationship between Enotourism supply and a winery's digital visibility represents a substantial gap in the extant literature.

To address this empirical lacuna, Baixauli-Gimenez et al. (2025) assess the role of Enotourism as a catalyst of online popularity. Their findings confirm that the adoption of enotourism activities generates a statistically significant increase in both web traffic and review volume across search platforms. Furthermore, the study reveals an asymmetry in this positive effect, with its magnitude being substantially larger among large-scale wineries.

3. Methodology

3.1. Data

For this study, a cross-sectional dataset was constructed comprising 4,348 observations corresponding to active Spanish wineries for the 2024 fiscal year. The sample was extracted from the Sistema de Análisis de Balances Ibéricos (SABI) (Bureau van Dijk, 2024), applying the CNAE code 1102 (Instituto Nacional de Estadística [INE], 2007) for wine production as the inclusion criterion.

Once the financial data were retrieved, 1,412 observations were removed due to complete missingness in competitive variables, and an additional 1,176 were excluded for exceeding a 20% missing-value threshold. Derived indicators such as value added per employee and revenue per employee were then computed. Subsequently, a Principal Component Analysis (PCA) with 3 components was performed to detect outliers. The PCA results revealed that, according to Hotelling's T^2 statistic, 43 anomalous observations were identified, with the F-value exceeding the 99th percentile (as illustrated in Figure 1). An additional 10 observations were flagged based

Article in Press

on the distance-to-model criterion, as shown in Figure 2, and were already present in the set of observations flagged by Hotelling's T^2 test.

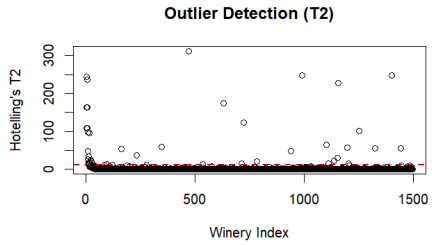


Figure 1. Hotelling's T2 outlier detection. Source: Own elaboration.

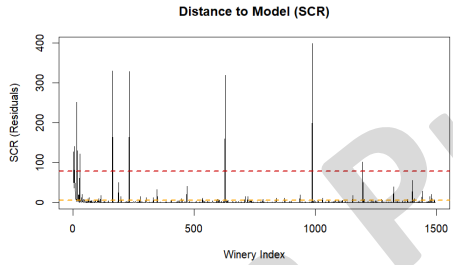


Figure 2. Distance to model (SCR). Source: Own elaboration.

To examine the impact of these observations, a descriptive analysis was conducted (Tables 1 and 2) comparing wineries identified as outliers with the remaining sample. The results indicate that these observations exhibit a substantially larger operational scale, with mean values for Revenue, Employees, Value Added (VA), and Assets that are considerably higher than those observed in the core dataset. This scale differential generates extreme dispersion across several financial indicators, particularly profitability measures such as ROA and ROE, as well as productivity indicators such as VA_Empl and Rev_Empl. The magnitude of these differences suggests that the identified wineries operate at fundamentally different scales, leading to disproportionate variability and distorting the distribution of key performance ratios. Consequently, these observations were excluded from the dataset prior to the econometric analysis.

Table 1. Descriptive analysis for outlier observations Source: Own elaboration

Variable	Mean	SD	Median	Min	Max
Revenue	43300.29	61816.65	5375.57	23.91	234035.26
Employees	128.53	228.83	12.00	1.00	937.00
VA	14090.23	19947.74	4415.53	-3444.89	78857.00

Article in Press

Assets	98527.21	132761.14	25235.70	54.13	447041.11
Profit	4513.54	8382.94	645.37	-4670.28	34629.00
ROA	-6.27	34.52	2.86	-137.06	58.45
ROE	-103.26	1111.70	8.24	-3668.05	3377.39
VA_Empl	300.86	859.38	90.63	-828.97	4415.53
Rev_Empl	544.30	934.73	286.83	3.80	5375.57

Table 2. Descriptive analysis for non-outlier observations Source: Own elaboration

Variable	Mean	SD	Median	Min	Max
Revenue	2243.92	5113.53	558.61	0.15	54883.61
Employees	9.98	14.68	5.00	1.00	159.00
VA	706.51	1617.42	200.64	-1993.44	16939.00
Assets	5751.88	11897.74	1796.72	15.04	134976.55
Profit	110.28	655.38	9.90	-3008.66	8057.99
ROA	1.96	8.50	1.99	-57.93	39.75
ROE	3.80	44.64	3.01	-576.16	447.35
VA_Empl	56.17	62.12	47.28	-177.46	590.98
Rev_Empl	183.72	198.06	127.17	0.05	2237.86

For winery-level web variables, data from Baixauli-Gimenez et al. (2025) were used to address the typical temporal lag in economic studies. The enotourism variable is binary, indicating whether a winery offers this activity. It was constructed by computing cosine similarity between the embedding of "the winery offers enotourism experiences" and each web text chunk. The highest-similarity chunk was processed via GPT-3.5 Turbo, and its output was classified against two reference embeddings using cosine similarity. After removing unmatched observations, the final sample comprised 872 wineries.

Additionally, from Baixauli-Gimenez et al. (2025), variables on digital marketing performance and web traffic were incorporated. These include estimated web visit volume via SimilarWeb, Google review count and average rating, and a binary indicator of TripAdvisor presence. Data extraction was automated in Python using Selenium WebDriver, with missing values imputed as zeros and interpreted as no activity. Finally, to capture environmental context, 2024 traveller and overnight stay statistics for the nearest tourist area were drawn from the Spanish National Statistics Institute (INE).

3.2. Models

Initially, the estimation of effects was conducted using a Difference-in-Differences (DiD) model, with the adoption of enotourism as the treatment variable. This methodology allows for isolating causal effects by comparing the temporal evolution of the outcome variable between a treated and a control group (Card & Krueger, 1994). However, the robustness and consistency

Article in Press

of this estimator require the fulfilment of a set of restrictive assumptions (Goodman-Bacon, 2021):

- Parallel trends: in the absence of the intervention, the evolution of the outcome variables across the treatment and control groups would have followed an identical trajectory.
- No anticipation: the units of analysis do not modify their behaviour in anticipation of the treatment prior to its effective implementation.
- Exogeneity: the event under analysis affects the variable of interest in isolation, without altering other covariates that may exert influence over the dependent variable.
- Stable Unit Treatment Value Assumption (SUTVA): the assignment of treatment to one unit does not affect the outcomes of other units, thereby ensuring the absence of spillover effects.

Within the geographical and commercial context of wineries, proximity effects and network externalities make it difficult to sustain compliance with SUTVA, as neighbouring wineries may exert mutual influence through shared tourism ecosystems and wine routes. Furthermore, adopting enotourism is a strategic decision that may be correlated with unobserved determinants of firm performance, raising concerns about endogenous treatment selection. Combined with the difficulty of identifying a structurally comparable control group, these limitations undermine the credibility of a DiD design in this setting.

Subsequently, an Instrumental Variables (IV) model was adopted, as this framework mitigates the endogeneity bias arising from omitted variables or simultaneity. Specifically, a winery's decision to integrate enotourism activities may be correlated with unobservable factors, such as managerial capacity or locational advantages, that also influence its Value Added (VA) (Wooldridge, 2010).

To isolate the pure causal effect, a Two-Stage Least Squares (2SLS) model was estimated. The first stage isolates the exogenous variation in enotourism by projecting it onto the instruments and the control covariates (winery age and size):

$$WineTourism_i = \gamma_0 + \gamma_1 Age_i + \gamma_2 Size_i + \gamma_3 SimilarVisits_i + \gamma_4 Trip_i + \gamma_5 Tourists_i + \nu_i \quad (1)$$

In the second stage, the fitted value of enotourism is employed to estimate its unbiased impact on Value Added, log-transformed to normalize financial dispersion:

$$VA_i = \beta_0 + \beta_1 Age_i + \beta_2 Size_i + \beta_3 \widehat{WineTourism}_i + \epsilon_i \quad (2)$$

The use of Similar Visits, Trip, and Tourists is justified on the grounds that they satisfy the relevance and exogeneity assumptions. While Similar Visits and Tourists capture aggregate demand fluctuations and global macroeconomic trends exogenous to local disturbances, Trip

Article in Press

introduces variation grounded in strictly external administrative and regulatory requirements. Collectively, these instruments predict the endogenous variable without correlating with the error term, thereby isolating exogenous variation and yielding unbiased estimates. (Wooldridge, 2010).

4. Results

The empirical models were implemented in the statistical software R using the `fixest` package, which provides efficient routines for instrumental variable estimation and econometric modelling (Bergé, 2018). The use of an open and script-based computational environment ensures full transparency, reproducibility, and replicability of the estimation procedure. All results reported below correspond to the Two-Stage Least Squares specification described in the previous section and are obtained from the final sample of 688 wineries.

4.1. First stage

To ensure robust estimation and address potential sources of bias, the adequacy of the 2SLS model was empirically assessed using three diagnostic tests. First, the Wu–Hausman test rejects the null hypothesis that the enotourism variable is exogenous (statistic = 4.765, p = 0.029), providing evidence of endogeneity and justifying the use of an instrumental variables estimator instead of Ordinary Least Squares (Wooldridge, 2010). Second, the relevance of the proposed instruments (Similar Visits, Trip, Tourists) was assessed through the first-stage regression. The corresponding F-statistic is 18.736 ($p < 0.001$), which exceeds the conventional threshold of 10 suggested in the econometric literature for detecting weak instruments (Staiger & Stock, 1997). This result indicates that the instruments are sufficiently correlated with the endogenous variable and that weak instrument concerns are unlikely. Table 3 reports the complete first-stage estimation results.

Table 3. First-stage regression results (Dependent Variable: Enotourism) Source: Own elaboration

ExplanatoryVariable	Estimate	Standard Error	t-value	
Intercept	0.255	0.197	1.291	
Similar Visits	0.000007	0.000003	2.380	*
Trip	0.340	0.050	6.800	***
Tourists	0.018	0.012	1.530	
Size	-0.008	0.015	-0.536	
Age	-0.001	0.001	-0.686	

Note: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

Finally, instrument validity was assessed through the Sargan overidentification test. The resulting statistic (2.990, p = 0.224) fails to reject the null hypothesis, suggesting that the instruments are jointly valid and correctly excluded from the structural equation (Sargan, 1958).

Article in Press

4.2. Second stage

The second-stage estimation results are reported in Table 4. The specification yields an adjusted coefficient of determination (adjusted R^2) of 0.692. Although goodness-of-fit measures have a limited interpretation in instrumental variable models, this value suggests that the included covariates capture a substantial portion of the variation in Value Added.

Table 4. 2SLS Estimation Results (Dependent Variable: VA) Source: Own elaboration

Explanatory Variable	Estimate	Standard Error	t-value
Intercept	-1.944	0.225	-8.650
Enotourism (Fitted)	0.644	0.250	2.574
Size	0.906	0.027	33.117
Age	0.007	0.003	-2.455

Note: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

The coefficient associated with the fitted value of Enotourism is positive and statistically significant ($\beta = 0.644$, $p < 0.05$), indicating that the adoption of enotourism activities is associated with higher levels of Value Added after accounting for potential endogeneity. Given that the dependent variable is specified in logarithmic form, this estimate implies that wineries offering enotourism activities exhibit, on average, substantially higher levels of Value Added compared to those that do not, ceteris paribus.

The estimated coefficient for Size is positive and statistically significant ($\beta = 0.906$, $p < 0.001$), indicating a positive relationship between firm size and Value Added. Firm Age also presents a positive coefficient that is statistically significant at conventional levels ($\beta = 0.007$, $p < 0.05$).

5. Conclusion

This study contributes to the literature by estimating the causal impact of enotourism on wineries' financial performance. Unlike conventional empirical approaches that assume exogeneity, the provision of tourism services is treated here as an endogenous strategic decision and addressed using a Two-Stage Least Squares Instrumental Variable (2SLS) framework. Diagnostic tests support the validity of the instruments and the suitability of the econometric specification to address potential self-selection bias.

The results indicate that the adoption of enotourism is positively and statistically significantly associated with higher levels of value added among wine-producing firms. In addition, firm characteristics, particularly operational scale, are also significantly related to economic performance. Overall, the analysis suggests that enotourism represents not only a marketing activity but also a relevant component of value creation in the wine industry.

Article in Press

References

Angrist, J. D., & Pischke, J. S. (2008). Mostly harmless econometrics: An empiricist's companion. Princeton University Press.

Baixauli-Gimenez, J., Debón, A., Cervelló-Royo, R., & Domenech, J. (2025). Turning tastes into trends: The power of enotourism in boosting wineries' popularity. Applied Economics, 1–18. https://doi.org/10.1080/00036846.2025.2600024

Bazen, S., Cardebat, J.-M., & Dubois, M. (2024). The role of customer and expert ratings in a hedonic analysis of French red wine prices: From gurus to geeks? Applied Economics, 56(46), 5513–5529. https://doi.org/10.1080/00036846.2023.2257036

Berge L (2018). Efficient estimation of maximum likelihood models with multiple fixed-effects: the R package FENmlm. CREA Discussion Papers, 13.

Bruwer, J., Coode, M., Saliba, A., & Herbst, F. (2013). Wine tourism experience effects of the tasting room on consumer brand loyalty. Tourism Analysis, 18(4), 399–414. https://doi.org/10.3727/108354213X13736372325957

Bureau van Dijk. 2024. "SABI: Sistema de análisis de balances ibéricos." https://sabi.bvdinfo.com/. Accessed: December, 2025.

Card, D., & Krueger, A. B. (1994). Minimum wages and employment: A case study of the fast-food industry in New Jersey and Pennsylvania. The American Economic Review, 84(4), 772–793.

Frost, W., Frost, J., Strickland, P., & Smith Maguire, J. (2020). Seeking a competitive advantage in wine tourism: Heritage and storytelling at the cellar-door. International Journal of Hospitality Management, 87, 102460. https://doi.org/10.1016/j.ijhm.2020.102460

Gómez, M., Pratt, M. A., & Molina, A. (2019). Wine tourism research: A systematic review of 20 vintages from 1995 to 2014. Current Issues in Tourism, 22(18), 2211–2249. https://doi.org/10.1080/13683500.2018.1441267

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. Journal of Econometrics, 225(2), 254–277. https://doi.org/10.1016/j.jeconom.2021.03.014

Instituto Nacional de Estadística [INE]. (2007). Real Decreto 475/2007, de 13 de abril, por el que se aprueba la Clasificación Nacional de Actividades Económicas 2009 (CNAE-2009). Boletín Oficial del Estado, 102, 18561–18562.

Sargan, J. D. (1958). The estimation of economic relationships using instrumental variables. Econometrica, 26(3), 393–415. https://doi.org/10.2307/1907607

Staiger, D., & Stock, J. H. (1997). Instrumental variables regression with weak instruments. Econometrica, 65(3), 557–586. https://doi.org/10.2307/2171753

Winfree, J., McIntosh, C., & Nadreau, T. (2018). An economic model of wineries and enotourism. Wine Economics and Policy, 7(2), 88–93. https://doi.org/10.1016/j.wep.2018.06.001

Wooldridge, J. M. (2010). Econometric analysis of cross section and panel data (2nd ed.). MIT Press.

Article in Press