Received 11 July 2022
Revised 6 November 2022
8 December 2022
Accepted 12 December 2022

Is "justice hurried actually justice buried"? An organisational perspective of the Italian criminal justice

Roberta Troisi and Gaetano Alfano
Department of Political and Communication Science, University of Salerno, Salerno, Italy

Abstract

Purpose – This study investigates the presence of a productivity–quality trade-off in judicial decisions from an organisational standpoint, focusing on the courts as bureaucracies. Applied to the Italian context and focusing on criminal courts, the main question addressed is whether or not increasing productivity diminishes decision quality.

Design/methodology/approach – Directional distance function (DDF) models were utilised to assess productivity. Two-sample *t*-tests are then used to compare the quality of efficient and inefficient units in first instance and appeal, with the aim to determine whether a productivity–quality trade-off exists.

Findings – The study's approach yields results that differ from previous studies. (1) The Italian judicial system is found less efficient. (2) The efficiency of the courts of first instance is relatively uniform. In contrast, there is a difference in efficiency between northern and southern courts of appeal, with northern courts on average being more efficient. (3) The analysis reveals a statistically significant productivity–quality trade-off when the courts of appeal are considered.

Research limitations/implications – New evidence of a judicial system is presented, suggesting reforms regarding "reasonable time" as the optimal balance between quality and productivity.

Originality/value – The organisational framework leads to evaluating the efficiency of the courts by considering the various types of proceedings based on the gravity/complexity of the cases. In light of the pyramidal structure of the justice system, the quality is then defined in terms of hierarchical control expressed as review rate.

Keywords Efficiency–quality trade-off, Organizational framework, Criminal justice, Bureaucracy, Proceedings, Hierarchical control, Reasonable time

Paper type Research paper

1. Introduction

Over the last ten years, there has been a significant increase in academic interest in judicial systems. It is not surprising, given that a functioning judicial system encourages business investment, job creation, institutional trust and economic growth (Zak and Knack, 2001; Tiede, 2018). On the contrary, it has been demonstrated that a dysfunctional legal system can have a negative impact on many aspects of development, including the legitimacy of the political system as a whole (Marciano *et al.*, 2019; Voigt, 2016). The length of the proceeding is the most investigated dimension of judicial efficiency. It is primarily understood as productivity, which is expressed as the number of cases resolved per capita in a given time while minimising the resources required (Voigt, 2016). This measure is fully evident and quantifiably accurate. It is also connected to the concept of "reasonable time" for judgements, which is the most frequently invoked element of a fair trial in international [1] and national

International Journal of Public Sector Management
Vol. 36 No. 1, 2023
pp. 94-109
Emerald Publishing Limited
0951-3558
DOI 10.1108/IJPSM-07-2022-0159

legal initiatives, as it is an essential guarantee that the demand of justice could be met in a timely manner, depending on the nature of the case [2].

Furthermore, studies on efficiency are frequently focused on the selection of the model analysis. These are designed to determine the optimal distribution of resources, typically focusing on a single court level and viewing it as just one stage in a larger proceeding with a three-tiered structure. However, measuring efficiency solely based on productivity leaves out a crucial part of how the legal system should work. Quality is equally a key dimension of the judicial efficiency as it relates to the reliability of the decisions and, consequently, is a major factor in fostering public confidence in the legal system. Efficiency and quality are complementary, resulting in a delicate equilibrium in which an increase in the number of cases may have negative effects on the quality, especially in terms of inaccuracy (Dimitrova-Grajzl *et al.*, 2016).

Since increased productivity cannot be achieved at the expense of quality, a quality assessment should accompany the efficiency assessment. Thus, the objective of this work is to examine the efficiency and quality of judicial systems from a new perspective, covering on the existing gaps. As a starting point, we propose an organisational framework that can address the research objectives, analytical methods and results. It views courts as bureaucratic organisations incorporated within a larger, multi-tiered judicial system (Canes-Wrone, 2003). Their organisational structure is founded on the criteria of balancing autonomy and control, which supports its overall functioning (Raine, 2000).

Autonomy is a horizontal organisational criterion that leads to a separation of the work of the courts, based on a different numerical composition, different roles and different activity, even though the decision-making activity is identical. This sharp division of labour in terms of activities, roles and competencies is reflected in the variety of proceedings used, characterised by increasing complexity, in relation to the gravity of the crime. Thus, a more serious crime carries harsher penalties, which are counterbalanced by enhancing the defendant's procedural guarantees, including major procedural steps than those established for less serious crimes, additional steps articulated in accordance with the dynamics of the proceedings, a larger number of judges for each step and finally the collegial composition of the court (see Section 2.1). Furthermore, the gravity of the crime frequently makes it difficult to obtain evidence. Particularly, demanding technical investigations or the difficulty of obtaining reliable testimony tend to lengthen the time required for each procedural stage (Coscas-Williams and Alberstein, 2019).

In contrast, control acts as a vertical organisational criterion defining the function of the higher judicial levels in order to maintain a balance between the lower courts autonomy and community interests. It is a form of bureaucratic control that permits the hierarchical review of judgements across multiple tiers of the judicial system, ensuring greater accuracy and reliability for the community. However, it has no direct effect on the lower court's judges, thereby preserving their autonomy. These organisational tenets direct the analysis while also allowing previous gaps to be filled. First, the judicial system is considered a complex hierarchy to assess as more than one single tier. Second, the productivity of courts is measured alongside the decisions of the first instance and courts of appeal by considering the different proceedings according to the complexity of the cases.

Third, by taking into account the pyramidal structure of the justice system, where the intermediate appeal level reviews a subset of trial decisions and the higher-level reviews a subset of appeal decisions, quality is expressed in terms of control, equally considered for both the first instance and courts of appeal. The decisions of the Supreme Court cannot be evaluated in terms of productivity or quality because the court is unique and incomparable to other courts and because its decisions are final and not subject to further review. Consideration of the third tier is limited to its hierarchical function of reviewing appeals decisions.

Italy is an interesting case when studying the judiciary. It appears that the sole evaluable aspect of the Italian judiciary involves performance-related factors. The Italian judiciary has

little interest in participatory reporting. Despite calls from European institutions for increased transparency and communication on the part of the Italian judiciary, particularly with citizens, there are currently few exemplary cases in this regard (Ricci and Pavone, 2020a).

In general, the performance of the Italian justice system is well below the European average. The European Court of Human Rights routinely penalises Italy for disregarding the "reasonable time" requirement for reaching a decision. Italy was among the six worst nations in the European Council for the number of court cases still pending, the low clearance rates and the lengthening of the disposition time (OECD, 2016).

We focus on criminal proceedings: as a distinct area of analysis, it appears to be underdeveloped in the existing literature on the efficient functioning of judicial systems. However, compared to civil trials, efficiency and quality in criminal proceedings involve a more difficult balance between individual guarantees, the certainty of the decisions and reasonable time. This is because it deals with a punishment that could imply a limitation of personal freedom.

The trial office reform law (no. 134/2021) was recently enacted in Italy in an effort to improve the efficiency of processes, primarily by shortening them. Particularly, for a number of crimes, it is stated that the decisions of the second and third tiers must be rendered within three years, failing which the prosecution will be rendered impracticable and the trial will cease. However, it is too early to determine if this intervention was successful.

Thus, in this study it is examined the presence of a productivity–quality trade-off in judicial decisions. Applied to the Italian context and focusing on criminal courts, the questions addressed are: (1) Does improving productivity come at the cost of the quality of decisions? (2) Along this line, are there any differences between the activity of the first instance and appeal criminal courts?

The analysis starts with the evaluation of the productivity of the judicial system through a directional distance function (DDF) model introducing as a novel output the number of resolved cases differentiated by proceedings. In the second step of the analysis, two-sample *t*-tests are run for the lower two (first instance and appeal) of the three tiers of the criminal court system to determine whether there are any differences in quality between efficient and inefficient units.

The Italian judicial system was found to be less effective than in studies employing DEA models without an organisational approach (Nissi *et al.*, 2019; Castro and Guccio, 2018). In addition, when the results are grouped by geographic region, they demonstrate that the level of efficiency of the courts of first instance is fairly uniform. In contrast, there is a difference in efficiency between northern and southern courts of appeal, with northern courts on average being more efficient. Regarding the productivity–quality trade-off, the analysis uncovered no sources of trade-off for courts of first instance. The difference in quality between efficient and inefficient courts of appeal is statistically significant ($p = 0.019$). In other words, productive courts of appeal have a higher average review rate. This contradicts previous research on the topic, which has consistently concluded that there is no trade-off between judicial quality and productivity.

The paper is structured as follows: Section 2 reviews the current literature; Section 3 presents the model; Section 4 describes the data and the results; Section 5 is about the discussion and Section 6 concludes.

2. Review of current literature

In this section, we explain why an organisational approach to judicial activity is essential for addressing the efficiency and quality of court decisions. When considering the individual courts as bureaucracy and the judicial system as a hierarchy, it is possible to provide relevant responses to the question posed above.

2.1 Courts as bureaucracy

Courts have received increased attention over the past decade (Visser *et al.*, 2019; Giacalone *et al.*, 2020), with a focus on achieving efficiency, timeliness and managerial effectiveness, in addition to the beginning of measuring court operations.

In Mintzberg's seminal work (1989), they were categorised as professional bureaucracies, but little research has been conducted from this perspective since then. Accordingly, courts are government agencies that are heteronomous public professional bureaucracies. This structure derives historically from relevant professions that have taken on a bureaucratic form. Their inherent characteristics of bureaucracy can have a significant impact on the efficiency and the quality of the courts.

Recognising the equal autonomy of its members is a crucial organisational and political criterion, as it frees the judges from any internal or external interference in their decision-making. It first addresses the organisational architecture in horizontal manner, in terms of high degree of separation between working units.

Consequently, this separation ensures a higher level of specialisation in judicial decisions, which ought to result in greater accuracy. The political significance lies in the fact that ensuring the well-being of citizens and the community as a whole is significantly more important today than in the past (Feeley, 2017).

This is particularly true in criminal processes, where a judge must determine the plea and its degree in order to set a penalty. Focusing on the organisational configuration of Tribunals, it is drawn on autonomy as a horizontal micro-criterion that orients a sharp division of the labour (Buta, 2021). The courts are different and separate for compositions. Thus, such a high degree of autonomy is primarily intended to create the most favourable conditions under which the judge may decide impartially, *sine spe ac metu* (without fear or hope), but also to avoid any mutual influence among judges with different criminal roles. As consequence, autonomy leads to different roles, different numerical composition of the courts, as well as different powers (Choi *et al.*, 2013). The difference in proceedings applied synthetises this sharp division of the labour by defining the steps to be taken in terms of activities, roles and competencies that are increasingly related to the gravity of the crime. Let us consider the Italian system of criminal proceedings (Ricci and Pavone, 2020b). The courts have broad authority over felonies and misdemeanours. However, if the crime is punishable by less than ten years in prison, a single professional judge will decide the case (trial court in single composition). When the crime carries a sentence of more than 10 years in prison, the case is decided by a panel of three professional judges (trial court in panel composition). This category is further divided into ordinary and extraordinary proceedings.

Ordinary proceedings consist of three phases: preliminary investigations, a preliminary hearing and the trial. Special proceedings are intended as an alternative means of adjudication, where time and steps are significantly reduced due to the nature of the evidence gathered by prosecutors or the defendant's request for a plea bargain (Coscas-Williams and Alberstein, 2019) (Figure 1).

The conclusion that can be drawn from these arguments is that autonomy impedes an overall measure of judicial productivity that does not account for internal separations and differences. In contrast, the examination of productivity must highlight the various proceedings utilised because they summarise the diverse work configurations that judges' autonomy necessitates.

2.2 Judicial hierarchy

An essential element of the bureaucratic organisation of the courts is the peculiar hierarchical control of the judicial system. It may be considered the vertical organisational criterion that directly affects the quality of decisions. The hierarchical structure of the judiciary has been

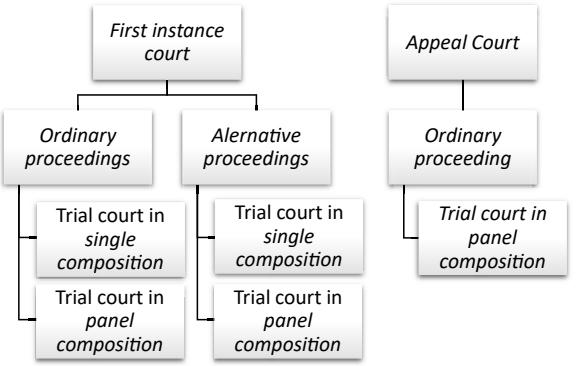

Figure 1.
The Italian system of criminal proceedings

studied as a "*sui generis*" control system over judges exclusively exercised through the review of the judicial decisions, without any direct consequences on the judges, to preserve their autonomy (Visser *et al.*, 2019). Following a regulation that is prevalent in public administration, where "the control" frequently has a three-tier level that are geographically characterised as local–districtual– national, the supreme authority has jurisdiction over the entire nation, and its words shall be final. Control in the judicial hierarchy is functional to achieving correctness, justification and finality principles in the decision rather than a traditional superior control activity on the lower judges (Halberstam, 2015). The purpose of a multi-tiered judicial system is to increase the overall quality of judicial outcomes. A second-stage decision should enable the elimination of any flaws that may have happened in the initial judgement, resulting in better and more trustworthy conclusions. In this regard, Cameron and Kornhauser (2009) argue that a judicial system organised on a three-tier hierarchy (trial, appeal and Supreme Court) as opposed to a two-tier one could demonstrate a higher degree of accuracy as the intermediate appeal level reviews a subset of trial decisions and the higher level reviews a subset of appeal decisions. In addition to the overarching objective of improving judicial decisions throughout the judicial levels, the rate of review by the superior court may be an essential measure of the quality of decisions made by lower courts.

Essentially, this is a vital productivity check. If an increase in productivity is coupled with an increase in review activity, the overall time gained in the lower grade is lost in the higher grade. Additionally, conflicting judgements from different levels have a tendency to impair the orienting capacity of the community and, consequently, its confidence (Tiede, 2018). The significance of the function of such a revision-based control provides two implications. The first is that efficiency and quality must be evaluated across all judicial tiers because it is inherent to the courts' organisational structure to operate at numerous levels. Any verification limited to a single grade is therefore insufficient. The second is decided by the review rate, which, despite certain limits, is a reasonably trustworthy indicator of judicial quality.

It could be argued that in the absence of methods of control, traditionally available to superiors within hierarchical organisations, avoidance of reviewing may provide a meaningful incentive for some judges in terms of reputation. Thus, a low rate of review may indicate the judge's ability to predict the preferences of higher courts (Choi *et al.*, 2013) rather than the better quality of the lower court's decision. Even with this bias, it should be preferred over the rate of appeal, a widely used criterion in the literature, although a manifestly imperfect proxy for quality. Litigants may appeal for trivial considerations of convenience, such as postponing judicial enforcement, in addition to genuinely observed errors in the lower court's decision (Dimitrova-Grajzl *et al.*, 2016). In this sense, the decision to

appeal for a second or third instance is endogenous to a losing litigant's estimated probabilities of success in the higher judicial tier (Priest and Klein, 1984).

The review rate may be a more trustworthy indicator of judicial quality than the appeal rate.

When reviewing a lower court's ruling, higher courts justify their interpretation by stating that lower courts "erred." There is a clear indication of the quality of a lower judicial decision. Even if a review of a decision by a higher court should not be deemed the correct disposition, a review is at least a new assessment of the facts or the law. It is a signal of differing points of view among the courts, which is damaging to both judicial certainty and public' perceptions of quality (Tiede, 2018).

The arguments of the aforementioned literature provide a theoretical framework through which the occurrence of a trade-off between productivity and quality of the judicial decisions synthesised will be assessed as follows.

(1) Criminal proceedings' efficiency takes into account the bureaucratic nature of the courts characterised by autonomy, synthetised in different kinds of proceedings with growing complexity according to the gravity of the crime.

(2) Criminal proceedings' quality is strictly tied to the hierarchical features of the judicial system. The indirect measure of the quality is the review rate of the decisions since it is more respectful of the role of the control that higher courts exercise on the lower courts. Low quality is when the higher courts rectify the lower court decision. Furthermore, there may be a perceived low quality in the discrepancy between the higher and lower courts. The same arguments can be used for high quality.

(3) As a multiple-tier system, productivity and quality are considered as related to the first two levels. This gives a more consistent idea of an overall functioning rather than focusing on one level.

3. The model

In this section, the analytical models used to analyse the productivity of the judicial system and then to investigate the presence of a productivity–quality trade-off in criminal judicial decisions in Italy are discussed.

Directional Distance Function (DDF) models were utilised to assess productivity. Falavigna and Ippoliti (2021) utilised it to assess the efficiency of the Italian judiciary. It is a variant of the traditional data envelope analysis (DEA), which is commonly employed to evaluate the efficiency of the legal systems (Yeung, 2018; Nissi *et al.*, 2019). DEA and DDF both consents to evaluate the efficient decision-making unit (DMU), also known as the frontier unit. For a given amount of input, this unit defines the maximum amount of output possible. Thus, efficiency has to be considered in terms of productivity, intended as an efficient quantity of judicial decisions on an annual basis (Fauvrelle and Almeida, 2018). In addition to the classic DEA model, DDF also agrees to take into account any undesirable outputs that are unavoidably linked to the production of an output (e.g. unprosecutable cases statute-barred by limitation). In particular, it defines the DMU as frontier units when, given a certain quantity of input, they maximise the desirable output while concurrently minimising the undesirable output. In cases where an undesirable output is inextricably linked to the production process, a standard DEA may lead to biased performance assessment (Färe *et al.*, 2007; Cooper *et al.*, 2007). The linear model that must be solved in order to determine efficiency is extensively described in Falavigna and Ippoliti (2021) and Álvarez *et al.* (2016).

The models are widely used due to their advantages, with the first being the absence of the specification of a formal relationship between the inputs and outputs (Eslamzadeh

et al., 2022). In addition, they enable the analysis of small samples with multiple inputs and outputs (Troisi and Alfano, 2022a).

After evaluating the DMU's productivity, the next step of the analysis is to determine whether a productivity–quality trade-off exists. Two-sample *t*-tests are used to compare the quality of efficient and inefficient units in first instance and appeal (Dimitrova-Grajzl *et al.*, 2016).

4. Data and results

This section describes the data sources, the variable utilised for the efficiency analysis and the quality variable employed for the productivity–quality trade-off evaluation.

4.1 Data sources

The primary data source was the website of the Italian Ministry of Justice. It gives information on the organisation of the Italian 26 judicial criminal districts, including the number of judges of first instance and appeal and their career advancement. In addition, it provides data regarding the number of cases handled in first instance and on appeal. For the first instance, it distinguishes between cases that are settled through an ordinary proceedings or a special proceeding (See Section 2.1). As for the appeal decisions, it distinguishes between those that confirm and those that review the first instance decision. The website of the Court of Cassation (http://www.italgiure.giustizia.it/sncass/) was accessed in order to gather additional data for the productivity–quality study. The Court of Cassation's decisions are provided in an aggregated format. In order to collect the necessary data for the quality proxy, 5,200 randomly selected decisions from the website of the Court of Cassation were consulted. The sample represents approximately 10% of the annual number of Supreme Court cases examined. The analysis focuses on the activities of Italian criminal courts in 2019. We chose not to consider more recent years because of the lockdown impact during the Covid-19 pandemic (2020–2021).

4.2 DDF variables

The selection of input and output variables is crucial to the application of the linear programme described in the preceding section.

The number of criminal judges is considered an input. This input is used in several works as a proxy of the dimension of the court (Castro and Guccio, 2018; Falavigna and Ippoliti, 2021). Additionally, as an input variable, the number of career advancement has been utilised. It serves as a proxy for the seniority and competence of judges (Guarnieri, 2004). The selection of this variable is consistent with recent research on the evaluation of public administration performance suggesting to account for the professional features of the agents involved (Ricci and Civitillo (2018). The last input variable is the courts' workload (new cases plus pending cases), often utilised in DEA as a proxy for the judicial demand (Ippoliti and Tria, 2020).

In order to implement the model outlined in the previous section, desirable and undesirable outputs for criminal first instance and courts of appeal have been identified. The number of resolved cases is considered the desirable output. Numerous research on judicial efficiency (Voigt, 2016) regard resolved cases as the output; in particular, studies evaluating the overall efficiency of the court system distinguish between civil and criminal cases, which are regarded as distinct products.

As highlighted in the previous section, first instance cases differ for the proceeding employed (ordinary or special proceeding) and for the seriousness of the crime (assigned to a single judge or a panel of judges). For this reason, to take account of the difference among the cases for the first instance model we consider as output: the number of cases resolved through an ordinary trial proceeding by a single judge (1) the number of cases resolved through an ordinary trial proceeding by a panel composition of three judges (2) the number of cases

resolved through a special trial proceeding by a single judge (3) the number of cases resolved through a special trial proceeding by a panel composition of three judges.

We thus considered two models. As described in Section 2.2, cases resolved by separate proceedings are unique to the first instance; therefore, they should be counted as separate output only for first instance criminal courts. In contrast, criminal courts of appeal are homogeneous in terms of both panel composition and proceedings, so the DDF model only takes into account the total number of resolved cases.

In both models, we consider cases that are ineligible for prosecution due to the expiration of the statute of limitations. Theoretically, these types of cases are viewed as a subclass of efficiency metrics, constituting aspects that influence the overall efficiency balance (Young and Singer, 2013). Analytically, they are considered undesirable output since they represent an inevitable source of delay of the trial. Their implications are in terms of costs of the judicial system, both perceived, as they contribute to the erosion of public confidence in the justice system (Voigt, 2016; Troisi and Alfano, 2022b), and real, as the proceedings took places without a judgement being reached.

To properly specify the DDF model, the output variables must satisfy two axioms: null-jointness and weak disposability. The decision-making process is time-consuming. The unavoidable delays needed to reach a verdict result in a number of unprosecutable instances, demonstrating the null-jointness property of the variables. Concerning the weak disposability of the outputs, the axiom states that if an input can generate outputs (desirables and undesirables), it is possible to decrease these outputs according to a reduction factor (Färe *et al.*, 2007). In the case of legal proceedings, this assumption is likewise confirmed. The only way to reduce the amount of unprosecutable cases is to reduce the decision time. This reduction is only possible if judges focus their efforts on a fewer number of cases. Consequently, the number of decisions will also drop, confirming the weak disposability requirements (Falavigna and Ippoliti, 2021). At last, desirable outputs are considered freely disposable.

After defining the input and output variables, the number of observations was examined as a final step. Cooper *et al.* (2007) propose that the number of DMUs must exceed $3*(\#input + \#output)$. In both models, the number of DMU considered is sufficient.

Table 1 summarises inputs and outputs. The table shows how all the data are characterised by a great variability among districts, both in the dimension of the courts, with courts of the first instance predictably greater than those of appeal, and in the number of resolved cases. The unprosecutable cases rate for first instance decisions is significantly higher than the one related to the appeal decisions.

4.3 Productivity–quality trade-off assessment variable

This section defines the quality measure exploited in the productivity–quality trade-off analysis. As said, in this study quality is expressed in terms of review rate that is the ratio of reviewed decisions to the total number of decisions evaluated by a higher court.

In particular, we distinguish between the number of reviewed decisions that are partially or totally overturned from the higher courts, based on the assumption that a fully revised decision expresses lower levels of quality than a partially revised decision.

As previously explained, the Italian criminal justice system employs hierarchical control to improve the quality of decisions at each level. Every type of review (whether factual or legal) is a symptom of the presumed low quality of the lower court's decision, with the objective difference between two decisions on the same issue increasing in proportion to the type of review (Halberstam, 2015). This measure reflects the characteristics of the Italian judicial system more accurately than other measures available in the literature, which typically focuses either on the reversal rate (Dimitrova-Grajzl *et al.*, 2016) or the overturned rate (Rosales-López, 2008; Mitsopoulos and Pelagidis, 2010).

	Mean (SD)	Min	Max	Model specification
Input variables				
First instance judges	168.38 (141.74)	24	573	Model 1st
Career advancement (first Instance)	8.00 (5.37)	2	19	Model 1st
First instance workload	43,942 (33,392)	7,202	152,973	Model 1st
Appeal judges	50.54 (41.24)	11	171	Model 2nd
Career advancement (appeal)	4.19 (2.72)	1	10	Model 2nd
Appeal workload	13,739 (16,201)	797	66,920	Model 2nd
Output variables				
Full trial resolved cases (first instance)				
Panel of judges (first instance)	346.84 (279.37)	50	1,072	Model 1st
Single judge (first instance)	6668.03 (4974.10)	1,332	21,391	Model 1st
Special proceeding resolved cases (first instance)				
Panel of judges (first instance)	105.23 (133.23)	4	634	Model 1st
Single judge (first instance)	3058.69 (2433.36)	189	9,474	Model 1st
First instance unprosecutable cases	12,034.42 (8798.31)	1848	34,266	Model 1st
Appeal resolved cases	2776.62 (2004.74)	479	7,497	Model 2nd
Appeal unprosecutable cases	1,082.62 (1214.31)	7	4,228	Model 2nd

Table 1.
Input and output variable summary statistics

Table 2 demonstrates that the mean proportion of reviewed cases is greater for courts of first instance than for courts of appeal, for both partial and total review. Table 2 also shows that the majority of reviews are partial, with an average ratio of two partial reviews to one total review for first instance courts and three partial reviews to one total review for appeal courts.

Overall considered, the percentage difference could depend on the different powers in reviewing lower courts decisions.

4.4 Results

In this section, analysis results are presented.

Table 3 presents DDF results for the 26 criminal districts, whereas Table 4 presents summary statistics grouped by geographical location.

The mean score for first instance and courts of appeal is 0.30 and 0.26, respectively. This suggests that, on average, for court of first instance and court of appeal, respectively, a score of 30 and 26% lower is needed to achieve the frontier of efficiency (0% inefficiency).

This demonstrates that the Italian criminal justice system as a whole is less efficient than in studies where DEA models were applied to both civil and criminal justice and outputs were not differentiated by proceedings (Nissi *et al.*, 2019). In addition, undesirable outputs (Falavigna and Ippoliti, 2021) lead to an increase in inefficiency when compared to studies that only consider the number of judgements produced while ignoring the number of nonprosecutable cases.

Grouping by geographical location (Table 4) reveals that the level of efficiency of the courts of first instance is fairly even. On the contrary, there is a distinction in terms of efficiency, with the most efficient courts of appeal concentrated in northern Italy (an average score of 0.22 compared to 0.30 of the southern regions). These results are not in line with research on the Italian judicial efficiency which show that northern courts consistently outperform southern courts (Castro and Guccio, 2018; Nissi *et al.*, 2019; Falavigna and Ippoliti, 2021). The efficiency gap between the north and south was reduced by taking into account the difference in proceedings.

The increased use of ordinary and panel composition proceedings in the South, as shown in Table 5, is indicative of an increase in case complexity, which is consistent with longer proceedings.

	First instance review rate			Appeal review rate		
	Total review	Partial review	Partial + total review	Total review	Partial review	Partial + total review
Ancona	0.17	0.35	0.52	0.10	0.26	0.36
Bari	0.34	0.34	0.69	0.14	0.17	0.31
Bologna	0.09	0.30	0.40	0.11	0.24	0.35
Brescia	0.11	0.63	0.51	0.07	0.23	0.30
Cagliari	0.15	0.28	0.42	0.06	0.19	0.25
Caltanissetta	0.15	0.46	0.62	0.10	0.24	0.34
Campobasso	0.33	0.45	0.60	0.09	0.22	0.31
Catania	0.14	0.32	0.44	0.08	0.26	0.34
Catanzaro	0.14	0.67	0.81	0.07	0.25	0.32
Firenze	0.25	0.48	0.63	0.09	0.19	0.28
Genova	0.18	0.53	0.62	0.02	0.10	0.12
L'Aquila	0.15	0.52	0.67	0.04	0.23	0.27
Lecce	0.17	0.35	0.52	0.12	0.11	0.23
Messina	0.26	0.53	0.48	0.11	0.40	0.51
Milano	0.12	0.30	0.44	0.08	0.24	0.32
Napoli	0.26	0.43	0.69	0.08	0.24	0.32
Palermo	0.17	0.24	0.41	0.08	0.21	0.29
Perugia	0.30	0.38	0.68	0.09	0.13	0.22
Potenza	0.21	0.25	0.46	0.13	0.28	0.41
Reggio Calabria	0.18	0.32	0.49	0.08	0.23	0.31
Roma	0.15	0.33	0.49	0.08	0.19	0.27
Salerno	0.31	0.44	0.64	0.14	0.24	0.38
Torino	0.22	0.50	0.57	0.06	0.20	0.26
Trento	0.13	0.35	0.49	0.10	0.28	0.38
Trieste	0.26	0.47	0.53	0.06	0.15	0.21
Venezia	0.21	0.52	0.52	0.06	0.19	0.25

Table 2. Judicial districts reviewed rates

The findings of the productivity–quality trade-off are presented in Table 6. The final two columns in Table 6 display the outcomes of the two sample *t*-tests. Consistent with previous research (Rosales-López, 2008; Mitsopoulos and Pelagidis, 2010; Dimitrova-Grajzl *et al.*, 2016), the analysis did not identify any sources of a productivity–quality trade-off for the courts of first instance. In contrast, when courts of appeal are considered, the null hypothesis of equal means must be rejected, both for partial and total reviewed cases. This indicates that the quality gap between efficient and inefficient courts is statistically significant ($p < 0.05$). In particular, the data highlight how courts with greater efficiency have, on average, a higher review rate. Figures 2 and 3 highlight how efficient courts of appeal are frequently characterised by extremely high review rates and *vice versa*.

5. Discussion

In this part, following the arguments presented in the literature section, the outcomes of the analysis are discussed.

First, the evaluation of the criminal judicial efficiency accounting for the bureaucratic nature of the courts pointed out some interesting arguments. Taking into account the different court proceedings, according to the complexity of the cases, allowed for results based on autonomy and the division of labour among criminal courts. Whereas the bulk of analysis about the efficiency of the Italian judicial system highlights a gap between northern and southern regions (Castro and Guccio, 2018; Nissi *et al.*, 2019), by considering the proceedings we found

	First instance DDF score	Appeal DDF score
Ancona	0.365	0.000
Bari	0.000	0.515
Bologna	0.000	0.000
Brescia	0.350	0.312
Cagliari	0.000	0.435
Caltanissetta	0.326	0.000
Campobasso	0.840	0.000
Catania	0.250	0.726
Catanzaro	0.115	0.544
Firenze	0.260	0.265
Genova	0.000	0.538
L'Aquila	0.000	0.458
Lecce	0.875	0.125
Messina	0.450	0.000
Milano	0.000	0.000
Napoli	0.000	0.631
Palermo	0.560	0.000
Perugia	0.220	0.383
Potenza	0.350	0.465
Reggio Calabria	0.450	0.724
Roma	0.340	0.545
Salerno	0.412	0.000
Torino	0.560	0.628
Trento	0.000	0.000
Trieste	0.540	0.075
Venezia	0.580	0.204

Table 3.
DDF score of the 26 judicial districts

	First instance model Mean (SD)	Appeal model Mean (SD)
Tribunal total	0.30 (0.26)	0.29 (0.26)
North	0.32 (0.26)	0.22 (0.25)
Centre	0.30 (0.07)	0.30 (0.23)
South	0.30 (0.29)	0.33 (0.29)

Table 4.
DDF score summary statistics

	Single judge			Panel of judges		
	Full trial resolved cases	Special proceeding resolved cases	Full trial/ special proceeding ratio	Full trial resolved cases	Special proceeding resolved cases	Full trial/ special proceeding ratio
North	73,582	38,205	0.52	3,611	881	0.24
Centre	39,774	16,837	0.42	2,443	696	0.28
South	76,500	26,692	0.35	3,803	816	0.21
Total	189,856	81,734	0.43	9,857	2,393	0.24

Source(s): Authors' elaboration on Minister of Justice data

Table 5.
Case filing statistics, broken down by proceeding, for first instance courts

	t-test (p-value) $H_0 : \overline{X}_1 - \overline{X}_2 = 0$ $H_1 : \overline{X}_1 - \overline{X}_2 \neq 0$	t-test (p-value) $H_0 : \overline{X}_1 - \overline{X}_2 = 0$ $H_1 : \overline{X}_1 - \overline{X}_2 > 0$
First instance		
Total review	0.335	0.168
Partial review	0.363	0.185
Partial + total review	0.226	0.113
Courts of appeal		
Total review	0.038	0.019
Partial review	0.010	0.005
Partial + total review	0.003	0.002

Note(s): In the t-test notation, $\overline{X}_1$ is the mean of the quality index of the efficient districts and $\overline{X}_2$ is the mean of the quality index of the inefficient district
The Shapiro–Wilk test is performed for checking normality of the sample and variance F-test for the equality of variance

Table 6.
Productivity–quality analysis results

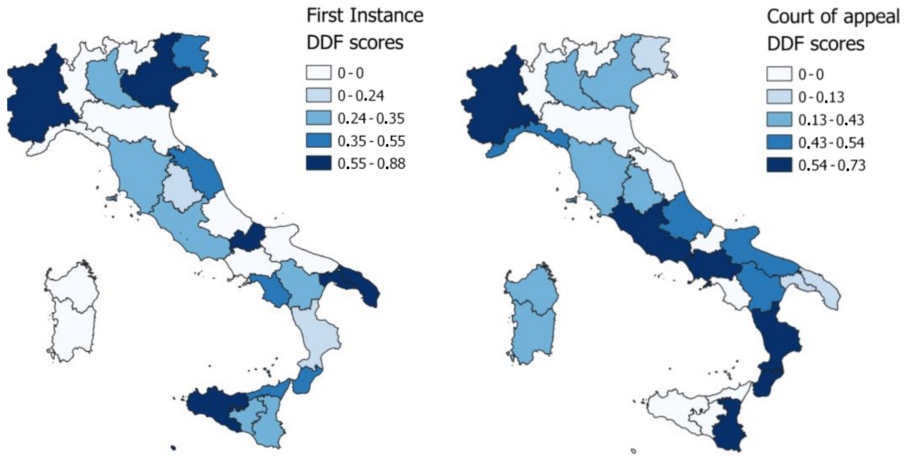

Figure 2.
Geographical illustration of the DDF scores

substantial homogeneity among the performances of the courts of first instance, with only the courts of appeal exhibiting geographical differences in terms of efficiency.

The implications go far beyond the merely methodological debate: productivity is multidimensional, strictly linked to the case complexity and consequently, the kind of proceedings. As shown in Table 5, ordinary and panel composition proceedings are utilised most frequently in the South, indicating a higher incidence of cases with greater complexity. Thus, it contributes to the lengthening of the criminal proceeding, which, if all cases are treated equally, results in an apparent decrease in production. In terms of public policy, adequate time is essential, especially in criminal proceedings. It has a direct bearing on the individual's right to obtain a final determination regarding the substantiated charge. However, the complexity and gravity of the cases must be taken into account when making timely decisions.

Second, the overall well-functioning of a judicial system is not determined solely by its productivity.

The quality of judicial decisions, expressed in this analysis as the hierarchical control of the superior tier, crucially complements the efficiency of the system and equally underpins the

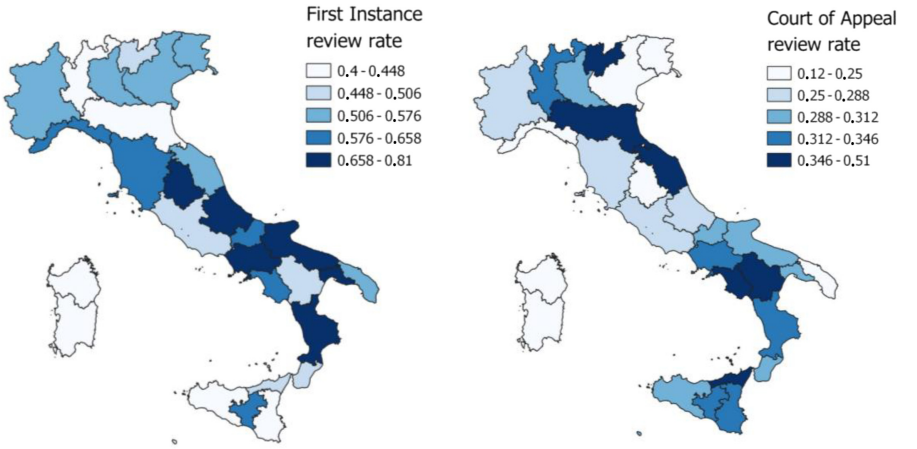

Figure 3.
Geographical illustration of the review rate (partial plus total review)

community's trust in the legal system (Troisi and Alfano, 2022b). The analysis revealed the existence of a productive–quality trade-off for the decision of the court of appeal. As for the efficiency, if the Court of Cassation reverses the decision of the court of appeal, the time saved will be countered or increased by a second appeal decision. As for the community's trust, an overturn of the decision of court of appeal would likewise be troublesome since it creates a discrepancy between decisions that cannot be further appealed. Two distinct decisions, even though the one of the higher tier naturally prevails, indicate two distinct orientations that guide the community with difficulty (Zak and Knack, 2001). Third, the judicial system as a complex bureaucratic system must be analysed as a whole, as any investigation of the efficiency and quality of a single tier may understate or amplify the presence of certain problems peculiar to that tier. For instance, the presence of the productivity–quality trade-off might have gone unnoticed if we had just investigated the first tier.

Partial examinations provide a representation of the judicial system that is not entirely accurate. Some studies, in confirming the existence of a North–South division solely based on the efficiency of a single tier, indicate the north of the country as a judicial context aligned with the best and most developed European countries and thus as optimal squares for foreign investment (Nissi *et al.*, 2019), missing the fact that the judicial system expresses quality and efficiency to be confirmed by its whole tiers.

6. Conclusions

This study investigates the presence of a productivity–quality trade-off in judicial decisions from an organisational standpoint, focusing on the courts as bureaucracies. Applied to the Italian criminal justice system, the two main questions are as follows: (1) Does increased efficiency diminish the quality of decisions? (2) Are there any distinctions between the activities of first instance and appeal criminal courts? According to the organisational perspective, first efficiency is assessed as a multidimensional measure of criminal court productivity, considering the different types of proceedings in relation to the severity/complexity of the cases.

Second, quality is tied to the hierarchical control of the judicial system and measured by the review rate throughout the three judicial levels.

Contrarily to previous studies, our analysis has highlighted the presence of a productivity–quality trade-off related to the courts of appeal decisions not instead shown for the courts of first instance decisions.

In light of these results, some considerations are necessary. For several reasons, this is a very significant finding. As it deals with the second tier of the judicial system, it is contrary to the idea of an increasing degree of accuracy along with the tiers. Moreover, it involves judges that are expected to have more experience than the first instance judges in balancing speed and accuracy. It has two crucial repercussions: as described, the reversed decisions reduce the efficiency because they lengthen the times that increased productivity had shortened. Furthermore, the overturned decisions on the part of the Supreme Court are final, so no further appealable. Lastly, if the citizens' perception of low quality could depend on a discrepancy between two decisions, the higher the courts are, the more it will be perceived as of low quality.

This study provides two key lessons useful at national and international level. In recent years, many European nations have enacted reforms aimed at increasing the productivity of the justice system while paying less attention to the quality of the judgements (CEPEJ, 2022). According to this approach, "reasonable time" was primarily recognised as a standard criterion for reducing time and thereby recognising a set of guarantees for people. This research has demonstrated that policies aimed at enhancing productivity must be carefully considered. At some point, an increase in quantity may be followed by a decrease in the quality of decisions. Furthermore, distinctive procedures associated with crimes have been shown to be connected to productivity.

If used correctly, the "reasonable time" criterion may provide the best balance of quality and efficiency. National and international reforms should create metrics of quick decision-making with nonuniform criteria that emphasise distinctions between crimes and related procedures. Furthermore, it should be evaluated as a full system performance rather than a single phase.

The following are the limitations of this study. The analysis was performed at the national level. The evaluation of the productivity–quality trade-off has international significance, particularly in the field of criminal justice, where it is an institution-based issue. However, the comparative perspective on judicial efficiency shows the limit to be affected by the diversity of legislations, whereas on the national level, factors that cannot be controlled by the courts (such as the legislation) but can affect their decisions are intrinsically controlled by the data. A further drawback of this work is the requirement to confirm the presence of the trade-off over a longer period of time (at least five years). Unfortunately, data for a more in-depth analysis were difficult to obtain. Future research could attempt to re-evaluate and broaden the analysis by building a judicial system panel dataset to examine any variations in judicial performance over time. Given the importance of the issue, it would be interesting to examine how the nature of the crime influences the trade-off and how it differs between the first and second instances.

Notes

1. Article 6ECHR.
2. E.g. Section 154, The German Code of Criminal Procedure; article 6, paragraph 2, Romanian Penal Procedure Code, art.542 Belgian code of criminal procedure.

References

Álvarez, I., Barbero, J. and Zofío, J.L. (2016), *A data envelopment analysis toolbox for Matlab (No. 2016/03). Universidad Autónoma de Madrid (Spain)*, Department of Economic Analysis (Economic Theory and Economic History), Washington.

Buta, B.O. (2021), "The autonomy of public defender's offices: a systematic comparison between Latin American countries", *International Journal of Public Sector Management*, Vol. 34 No. 5, doi: 10.1108/IJPSM-09-2020-0241.

Cameron, C.M. and Kornhauser, L.A. (2009), "Modeling collegial courts (3): adjudication equilibria", *NYU School of Law, Public Law Research Paper*, Vol. 12 No. 34, pp. 9-29, doi: 10.2139/ssrn.1400838.

Canes-Wrone, B. (2003), "Bureaucratic decisions and the composition of the lower Courts", *American Journal of Political Science*, Vol. 47 No. 2, pp. 205-214, doi: 10.1111/1540-5907.00014.

Castro, M.F. and Guccio, C. (2018), "Measuring potential efficiency gains from mergers of Italian first instance Courts through a nonparametric model", *Public Finance Review*, Vol. 46 No. 1, pp. 83-116, doi: 10.1177/1091142116652723.

CEPEJ (2022), available at: https://rm.coe.int/cepej-report-2020-22-e-web/1680a86279

Choi, S.J., Gulati, M. and Posner, E.A. (2013), "How well do measures of judicial ability predict judicial performance? A case study using securities class actions", *International Review of Law and Economics*, Vol. 33, pp. 37-53, doi: 10.1016/j.irle.2012.09.002.

Cooper, W.W., Seiford, L.M. and Tone, K. (2007), *Data Envelopment Analysis: A Comprehensive Textwith Models, Applications, References and DEA-Solver Software*, Princeton University Press, Princeton, NJ, ISBN: 978-0-306-47541-2.

Coscas-Williams, B. and Alberstein, M. (2019), "A Patchwork of doors: accelerated proceedings in continental criminal justice systems", *New Criminal Law Review*, Vol. 22 No. 4, pp. 585-617, doi: 10.1525/nclr.2019.22.4.585.

Dimitrova-Grajzl, V., Grajzl, P., Slavov, A. and Zajc, K. (2016), "Courts in a transition economy: case disposition and the quantity–quality trade-off in Bulgaria", *Economic Systems*, Vol. 40 No. 1, pp. 18-38, doi: 10.1016/j.ecosys.2015.09.002.

Eslamzadeh, S.M.K., Grilo, A., Espadinha-Cruz, P., Rodrigues, J.P.C. and Lopes, J.P. (2022), "A framework for fire departments' performance assessment: a systematic literature review", *International Journal of Public Sector Management*, Vol. 35 No. 6, doi: 10.1108/IJPSM-06-2021-0149.

Färe, R., Grosskopf, S. and Pasurka, C.A. Jr (2007), "Environmental production functions and environmental directional distance functions", *Energy*, Vol. 32 No. 7, pp. 1055-1066, doi: 10.1016/j.energy.2006.09.005.

Falavigna, G. and Ippoliti, R. (2021), "Reform policy to increase the judicial efficiency in Italy: the opportunity offered by EU post-Covid funds", *Journal of Policy Modeling*, Vol. 43 No. 5, pp. 923-943, doi: 10.1016/j.jpolmod.2021.06.001.

Fauvrelle, T.A. and Almeida, A.T.C. (2018), "Determinants of judicial efficiency change: evidence from Brazil", *Review of Law and Economics*, Vol. 14 No. 1, doi: 10.1515/rle-2017-0004.

Feeley, M. (2017), "Two models of the criminal justice system: an organizational perspective", in Hancock, B.W. and Sharp, P.M. (Eds), *Crime, Law and Society*, Prentice-Hall, pp. 119-137.

Giacalone, M., Nissi, E. and Cusatelli, C. (2020), "Dynamic efficiency evaluation of Italian judicial system using DEA based Malmquist productivity indexes", *Socio-Economic Planning Sciences*, Vol. 72, 100952, doi: 10.1016/j.seps.2020.100952.

Guarnieri, C. (2004), "Appointment and career of judges in continental Europe: the rise of judicial self-government", *Legal Studies*, Vol. 24 Nos 1-2, pp. 169-187, doi: 10.1111/j.1748-121X.2004.tb00246.x.

Halberstam, Y. (2015), "Trial and error: decision reversal and panel size in state courts", *The Journal of Law, Economics, and Organization*, Vol. 32 No. 1, pp. 94-118, doi: 10.1093/jleo/ewv007.

Ippoliti, R. and Tria, G. (2020), "Efficiency of judicial systems: model definition and output estimation", *Journal of Applied Economics*, Vol. 23 No. 1, pp. 385-408, doi: 10.1080/15140326.2020.1776977.

Marciano, A., Melcarne, A. and Ramello, G.B. (2019), "The economic importance of judicial institutions, their performance and the proper way to measure them", *Journal of Institutional Economics*, Vol. 15 No. 1, pp. 81-98, doi: 10.1017/S1744137418000292.

Mintzberg, H. (1989), "The structuring of organizations", *Readings in Strategic Management*, Palgrave, London, pp. 322-352.

Mitsopoulos, M. and Pelagidis, T. (2010), "Greek appeals Courts' quality analysis and performance", *European Journal of Law and Economics*, Vol. 30 No. 1, pp. 17-39, doi: 10.1007/s10657-009-9128-4.

Nissi, E., Giacalone, M. and Cusatelli, C. (2019), "The efficiency of the Italian judicial system: a two stage data envelopment analysis approach", *Social Indicators Research*, Vol. 146 No. 1, pp. 395-407, doi: 10.1007/s11205-018-1892-5.

OECD (2016), available at: https://www.oecd.org/gov/access-to-justice.htm

Priest, G.L. and Klein, B. (1984), "The selection of disputes for litigation", *The Journal of Legal Studies*, Vol. 13 No. 1, pp. 1-55, available at: https://www.jstor.org/stable/724341

Raine, J.W. (2000), "Modernising courts or courting modernisation?", *International Journal of Public Sector Management*, Vol. 13 No. 5, doi: 10.1108/09513550010350788.

Ricci, P. and Civitillo, R. (2018), "Italian public administration reform: what are the limits of financial performance measures?", *Outcome-Based Performance Management in the Public Sector*, Springer, Cham, pp. 121-140, doi: 10.1007/978-3-319-57018-1_7.

Ricci, P. and Pavone, P. (2020a), "Better justice towards sustainability: the case of an accountable judicial institution in Italy", *International Journal of Public Sector Performance Management*, Vol. 6. doi: 10.1504/IJPSPM.2020.10034492.

Ricci, P. and Pavone, P. (2020b), "The experience of social reporting in Italian judicial offices. The laboratory of the public prosecutor's office in Naples", *International Journal of Public Sector Management*, Vol. 33. doi: 10.1108/IJPSM-04-2020-0102.

Rosales-López, V. (2008), "Economics of Court performance: an empirical analysis", *European Journal of Law and Economics*, Vol. 25 No. 3, pp. 231-251, doi: 10.1007/s10657-008-9047-9.

Tiede, L.B. (2018), "The rule of law, institutions, and economic development", in May, C. and Winchester, A. (Eds), *Handbook on the Rule of Law*, Edward Elgar Publishing, pp. 405-418, doi: 10.4337/9781786432445.00034.

Troisi, R. and Alfano, G. (2022a), "Proximity and inter-firm corruption: a transaction cost approach", *Small Business Economics*. doi: 10.1007/s11187-022-00649-y.

Troisi, R. and Alfano, G. (2022b), "The re-election of corrupt mayors: context, relational leadership and level of corruption", *Local Government Studies*, pp. 1-22, doi: 10.1080/03003930.2022.2087060.

Visser, M., Schouteten, R. and Dikkers, J. (2019), "Controlling the courts: new public management and the Dutch judiciary", *Justice System Journal*, Vol. 40 No. 1, pp. 39-53, doi: 10.1080/0098261X.2018.1539645.

Voigt, S. (2016), "Determinants of judicial efficiency: a survey", *European Journal of Law and Economics*, Vol. 42 No. 2, pp. 183-208, doi: 10.1007/s10657-016-9531-6.

Yeung, L. (2018), "Measuring efficiency of Brazilian courts: one decade later", *Revista de direito administrativo*, Vol. 279, pp. 111-134, doi: 10.2139/ssrn.3200588.

Young, W.G. and Singer, J.M. (2013), "Bench presence: toward a more complete model of federal district court productivity", *Penn State Law Review 55*, Vol. 118 No. 1, pp. 55-98.

Zak, P.J. and Knack, S. (2001), "Trust and growth", *The Economic Journal*, Vol. 111 No. 470, pp. 295-321, doi: 10.1111/1468-0297.00609.

Corresponding author
Roberta Troisi can be contacted at: rtroisi@unisa.it