Pressure on judges: How managerialisation and evolving professional standards affect judges’ autonomy, efficiency and stress

In recent decades, many judicial systems have witnessed the “managerialisation” of justice—a phenomenon involving the application of private sector techniques to enhance court functioning. These techniques encompass engaging professional managers, reorganizing services, implementing performance measures, and adopting performance-based budgeting. Balancing these approaches with judges’ professional standards, as independence and quality, raises critical questions. How does managerialisation affect judges’ organization, autonomy, work quality, efficiency, and work-related stress? A survey conducted from June to December 2020 among first-instance judges in Finland, Italy, and the Netherlands, with distinct budgeting models, aimed to understand the impact. The questionnaire explored pressure sources, perceived stress levels, and mechanisms mitigating pressure and stress. Results showed significant differences in work organization, performance targets, and judge autonomy. However, similarities emerged in perceived pressure, work-related stress, and motivation. The study suggests that while not a panacea, managerialisation doesn’t inherently increase pressure and stress. It may improve court organization, clarifying the interplay between professional standards and financial considerations.

Key words

Performance management; court management; professional standards; judicial administration; judges’ autonomy; efficiency; court budgeting
Resumen

En las últimas décadas, muchos sistemas judiciales han sido testigos de la “gerencialización” de la justicia, un fenómeno que implica la aplicación de técnicas del sector privado para mejorar el funcionamiento de los tribunales. Estas técnicas abarcan la contratación de gestores profesionales, la reorganización de los servicios, la aplicación de medidas de rendimiento y la adopción de presupuestos basados en el rendimiento. Equilibrar estos enfoques con las normas profesionales de los jueces, como la independencia y la calidad, plantea cuestiones críticas. ¿Cómo afecta la gerencialización a la organización, la autonomía, la calidad del trabajo, la eficiencia y el estrés laboral de los jueces? Una encuesta realizada entre junio y diciembre de 2020 entre jueces de primera instancia de Finlandia, Italia y los Países Bajos, con distintos modelos presupuestarios, tenía como objetivo comprender el impacto. El cuestionario exploraba las fuentes de presión, los niveles de estrés percibidos y los mecanismos para mitigar la presión y el estrés. Los resultados mostraron diferencias significativas en la organización del trabajo, los objetivos de rendimiento y la autonomía de los jueces. Sin embargo, surgieron similitudes en la presión percibida, el estrés laboral y la motivación. El estudio insinúa que, aunque no es una panacea, la gerencialización no aumenta intrínsecamente la presión y el estrés. Puede mejorar la organización de los tribunales, aclarando la interacción entre las normas profesionales y las consideraciones financieras.

Palabras clave

Gestión del rendimiento; gestión de los tribunales; normas profesionales; administración judicial; autonomía de los jueces; eficiencia; presupuesto de los tribunales
Table of contents

1. Introduction ............................................................................................................................ 4
2. Theoretical framework: managers and professionals at the courts ................................. 5
3. Methodology ........................................................................................................................... 8
   3.1. Choice of countries ....................................................................................................... 9
   3.2. Definition of “pressure on judges” ........................................................................... 11
   3.3. Variables ..................................................................................................................... 11
   3.4. Statistical analysis ..................................................................................................... 14
4. Results .................................................................................................................................... 14
   4.1. General outcomes ..................................................................................................... 14
   4.2. Characteristics of respondents .................................................................................. 15
   4.3. Question 1: are judges working under pressure? .................................................. 17
   4.4. Question 2: where is the focus of management? .................................................... 21
   4.5. Question 3: does performance management exert pressure on judges? ............. 24
   4.6. Question 4: are professional standards discernible and do they interact with
        performance targets? ..................................................................................................... 26
   4.7. Question 5: is extrinsic and intrinsic motivation related to stress and what is the
        role of management in regulating stress? ..................................................................... 27
5. Differences among judges ................................................................................................... 29
6. Determinants of work-related stress ................................................................................. 30
7. Conclusions ........................................................................................................................... 31
References .................................................................................................................................. 33
Appendix 1. Content-related survey questions, organised by topic..................................... 37
Appendix 2. Results of ordinal regression analysis ............................................................... 39
1. Introduction

In the last decades, many judiciaries have been affected by a phenomenon that can be described as “managerisation”, meaning the application of private sector techniques to courts to improve the functioning of justice. The application of these techniques involves the engagement of professional managers in courts, the re-organisation and optimisation of services, the use of performance measures and performance management and the implementation of new budgeting models. This phenomenon is also known as New Court Management (NCM), referring to the application of New Public Management methods to courts a decade later than the rest of the public sector (Fabri and Langbroek 2000). As these methods have found their way into the courts, the role of management has increased, and the objectives of the courts have shifted. As Lienhard and Kettiger put it succinctly, “the prevailing opinion is that court management should primarily or exclusively serve to ensure (i) the effective protection of legal rights (in particular the right to a timely and objective decision based on a fair procedure) and (ii) the efficient expenditure of public funds.” (Lienhard and Kettiger 2017, p. 12). The first objective represents the typical view of judges, but the second elevates the importance of efficiency to a much higher level than before. Nobody will be against the efficient use of taxpayers’ money, but it was not, and is not, at the forefront of the minds of judges when it comes to adjudication, as our survey will show. Trying to achieve a reorientation is one thing, actually reaching it is another.

Many judges have misgivings about the benefits of managerialism (Holvast and Doornbos 2015, De Santis et al. 2016). They share the view of other professionals who blame management for diminishing the space for professionals to do their job properly, reducing professionals to bureaucrats and burdening them with administrative tasks (Noordegraaf and Steijn 2014). As to the latter, instruments such as performance budgeting require data about production processes, and these need to be registered. Also, the emphasis on efficiency leads, in their view, to insufficient budgets that are based on highly theoretical estimates of the possibilities for efficiency improvement. In Switzerland, judges argue that managerialism reduces the courts to a shoe factory (De Santis et al. 2016, p. 130) or, in the Netherlands, a biscuit factory (Frissen et al. 2014, Holvast and Doornbos 2015). These metaphors reflect a genuine fear of Taylorism. There is also deep resistance against the terminology of management, which is seen as not applicable to or appropriate for the judiciary. As De Santis et al. (2016, p. 130) note, concepts like productivity, products and customers are resisted. We may add production targets and the terminology of strategy development, such as mission and vision. Also, measurement and quantitative analysis are often regarded with suspicion, reflecting doubts about usefulness in relation to the effort it takes to gather data.

However, professionals are not passive victims of managers and bureaucrats (Newman 2013). They shape their professional practice. Among professionals, this holds true in particular for judges who are supposed to be independent and can take recourse to the courts themselves if they feel their independence is at stake. Judges that hold governments and multinational companies to account in their judgments can hardly be expected to bow to management easily. They regularly refuse to go along with human resource management techniques such as performance evaluation. Collectively, they can and actually do make their views known about the governance and management of the
judiciary to councils for the judiciary, politics and media, with a large impact. An example is the manifesto of Dutch judges of 2012 (Holvast and Doornbos 2015, Berendsen et al. 2015). Consequently, the management of a court is a complex task, especially because how far judicial independence reaches is arbitrary. Thus, next to stressed professionals, management, often made up of judges, is squeezed between external demands and constraints on resources, and the professional standards of judges.

To examine the interaction of professionals and management under different conditions of “managerialism” from the perspective of the judges, between June and December 2020, a survey was conducted among all first-instance judges in Finland, Italy and the Netherlands. In these countries three different management and budgeting models are in place (Viapiana 2018, 2019, 2020). The purpose of the survey was to investigate the level of pressure on judges and their work-related stress, and the sources of pressure. Workload, performance targets regarding production, timeliness and efficiency, organisational autonomy, motivation and other elements that can positively or negatively affect work well-being were investigated. The main question that we attempt to answer is how different management practices exert pressure on judges and impact their work-related stress.

Section 2 describes the theoretical framework, and Section 3 the methodology of the survey. Section 4 presents the results with a focus on the differences among countries. Section 5 examines the homogeneity of the response of the judges, while section 6 briefly examines the connections between the variables by means of an ordinal regression analysis with work-related stress as the dependent variable. Section 7 concludes with a discussion of the main findings.

2. Theoretical framework: managers and professionals at the courts

In the 1980s, a business-like approach called New Public Management (NPM) was developed to improve efficiency in the public sector by applying management techniques from the private sector (Hood 1991, Osborne and Gaebler 1992). The judiciary was initially slow to adopt NPM principles due to concerns about judicial independence and a non-managerial attitude within the legal profession (Maier 1999, Fabri and Langbroek 2000, Fabri et al. 2003). Traditionally, the evaluation of the functioning of the courts focused solely on legal aspects, disregarding factors such as time and cost (Contini and Carnevali 2010). This single focus does not mean that there were no quality issues. For instance, McBarnet argued in the eighties that in first instance courts in the UK legality and justice were compromised (McBarnet 1983). In the late 1990s, an increase in caseloads combined with budget cuts led to longer proceedings and a decline in public trust. Courts were seen as expensive, slow, and inefficient (Fabri et al. 2003, Contini and Mohr 2008, Bunjevac 2017). The “litigation explosion” and changing social expectations also highlighted the need for efficient and effective courts (Van Dijk 2014, Rizos et al. 2021). The globalized world further emphasized the importance of efficient courts, as their functioning affects the economy, markets and investments. Research has shown a connection between well-functioning judicial systems and economic growth. For example, the quality and efficiency of judicial systems can influence the size of firms, foreign direct investments and interest rates (see Rizos et al. 2021).
To meet the demand for greater efficiency and timeliness, judiciaries began adopting the NPM perspective in what became known as “New Court Management,” in the mid-1990s. This shift focused on individuals (citizens) rather than institutions, viewing justice as a public service rather than a source of power. Elements such as cost control, performance measurement, case clearance and reduced delays became important aspects of court management (Maier 1999, Fabri and Langbroek 2000, Fabri et al. 2003). The application of these techniques also affected judges, as performance measures and performance management became an integral part of court management.

To assess the potential impact of NCM on the performance of judges, it is necessary to consider what drives judges. According to Posner (2008), judges are like any other agent rational and self-interested individuals aiming at maximising their personal utility (“homo economicus”). “Personal utility” does not necessarily regard income. It can also be related to career, visibility, power, success, professional reputation and so on. A more recent stream of studies, however, has emphasised the importance of socio-psychological forces that motivate individuals. Many studies in the public and private sector have demonstrated that individuals are driven by material and non-material motivations such as social motivations, in particular striving for approval, imitation, group loyalty, adherence to norms and affective social ties (Akerlof 1983, Baumeister and Leary 1995, Fehr and Gächter 2000, Van Dijk et al. 2002, Van Winden 2015) and internal motivations such as moral beliefs and values (Deci and Ryan 1985, Minkler 2004). In most judiciaries judges are appointed until retirement, and for most of them it is indeed a lifetime employment. They are in it for the long run and make their career there. This strengthens the importance of motivations other than short-term financial gain, making judges a close-knit group with clearly defined norms, strong group loyalty and close social ties. They have strong incentives to fit in and be seen as competent judges by their peers, including the judges at the appeal courts and the supreme court. This results in strong intrinsic motivations, and, as these motivations tend to be homogeneous among judges, strong professional values and standards within the group of judges, centred around independence, impartiality and high legal quality. As Noordegraaf and Steijn (2014) argue about professionals in general, the more standards they set, the stronger they are, also with respect to autonomy and power.

Values and standards find a strong common foundation in supranational legal instruments. In Europe, Art 6 ECHR provides clear guidance: “In the determination of his civil rights and obligations or of any criminal charge against him, everyone is entitled to a fair and public hearing within a reasonable time by an independent and impartial tribunal established by law” (Council of Europe 1950). Independence extends obviously to the content of judgments but also to what the judge deems necessary to reach a fair decision within the confines of the law. The authority to determine the procedure and the time needed to adjudicate cases are part of independence. But independence may extend further to subjects as varied as the requirements for digital proceedings, including online hearings, and the content of professional training.

However, while independence is a paramount value, judges, as long as they have a career perspective, are to some degree dependent on the authorities that decide or advise on promotions (Schneider 2005, Robinson 2007). This results in incentives to perform in a way these functionaries value. Career considerations lead, in essence, to a “rank order
treatment” which has been shown to generate strong incentives to achieve the targets set by authorities (Bull et al. 1987, Van Dijk et al. 2001). In the courts, these functionaries are generally either the president of the court or a council of the judiciary. This dependence may result in judges being susceptible to the views of these functionaries. It may strengthen the hold of professional standards further, but, if the functionaries have an NCM orientation, it may also lead to the acceptance of other orientations, such as timeliness and efficiency. In line with these potentially conflicting values, promotions prove to be a sensitive area as the surveys by the European Network of Councils of the Judiciary (ENCJ) among European judges show. In many countries substantial percentages of judges believe that judges have been promoted for reasons other than on the basis of ability and experience (for instance, 36% in Spain and 24% in Germany).¹ To conclude, judges have strong internal motivations, but they are also subject to external incentives. NCM strengthens these external incentives.

What is to be expected from strengthening extrinsic (career) incentives in a profession that has a strong intrinsic motivation? In a simplistic version of NPM logic that disregards intrinsic motivation this will lead to higher productivity. According to Robinson and Brumby (2015, p. 51) the “process of making desired outcomes as explicit as possible, and linking output and activities to those outcomes can be a means of improving goal alignment.” Using this reasoning, performance budgeting, by making the link between performance and funding explicit, and performance management contribute to the understanding of the goals and, in this way, improve productivity. This, however, does not take into account the effects of the replacement of intrinsic motivation by extrinsic motivation, which can be counterproductive.² These effects can be particularly impactful in a situation where intrinsic motivation is strong as in the judiciary. Efforts by management to introduce extrinsic motivators, thereby unwittingly or unwittingly crowding out intrinsic motivation, are likely to meet resistance, and may lead to bureaucratisation and not higher efficiency. The expectation that judges will work harder by implementing control mechanisms follows bureaucratic logic and may well prove to be unrealistic among professionals in general, and judges in particular.

If the intended effect of extrinsic incentives is not so much an increase of overall productivity but a reorientation of objectives from a single focus on legal excellence to a broader notion of quality as valued by the court users, including timeliness and giving attention to cases proportional to their relevance (Frissen et al. 2014), the underlying reorientation can be seen as a response to genuine demands in society. Still, such a reorientation is likely to meet resistance as well. Part of NCM is the introduction of methods of strategy development in the courts to get judges thinking about the evolving needs of the population they serve. Under this new paradigm of court-user orientation, judges are no longer only independent decision-makers, but they are actively part of a public organisation delivering services to the public (Contini and Mohr 2008). Sharing such a new “mission” and “vision” of the organisation would help to align intrinsic and

¹ Percentage of respondents that agrees with the statement “I believe judges in my country have been promoted or appointed other than on the basis of ability and experience during the last three years” (ENCJ 2022, Table 21). For the countries participating in this survey: Italy 41%, Finland 11% and the Netherlands 5%.

² For a general discussion of the literature, see Gneezy et al. (2011).
extrinsic motivation within the judiciary. However, we already noted that judges resist strategy development.

Judges and management can interact in different ways, depending on the role management chooses. This leads to more or less pressure on judges and work-related stress. Tension and conflicts may arise in different degrees. In the Swiss study, already quoted, terms are used like two worlds and various logics, but it is also found that the views of judges and managers are not totally incompatible (De Santis et al. 2016, p. 128). The study shows that both groups share at least half of the expectations of what is, in their terms, a good judiciary and that these expectations are compatible with NPM. As to the differences, judges emphasise humane aspects of justice, fairness and impartiality, while management focuses on customer orientation, including efficiency, accessibility and timeliness. It should also be noted that differences between the two groups are smaller when management consists of judges. We conclude that the impact of the introduction of NPM on judges and their reactions to it are not clear-cut and are a matter of empirical research.

It should be emphasized that pressure on judges and work-related stress, affecting motivations, highly depend on management practices but also on the external constraints management is faced with. As summarised in the review by Casaleiro et al. (2021), there are several descriptive studies on working conditions, stress, and job satisfaction of judicial professionals. Some of these studies show that work-related stress is connected to an excessive volume of work and a high-speed work pace (Rogers et al. 1991, Ferreira et al. 2014) and the organisational structure that is largely determined by law (Ciocoiu et al. 2010). Another study by Na et al. (2018) suggests that organisational support compensates for work-related stress. These studies show a general dissatisfaction with working conditions and, in particular, the intensity of work, stemming from heavy caseload, backlogs and time constraints (Casaleiro 2021, p. 23). Much of these issues are related to management, but not exclusively, as budgetary constraints and legal constraints on the organisation of the judiciary and the design of judicial procedures also often play a large role. The mentioned studies have limitations, because they consider a single jurisdiction such as immigration disputes or a single judicial area (Casaleiro et al. 2021, p. 17). The present research adds to this literature by comparing pressure on judges and work-related stress in the judiciaries of three countries.

3. Methodology

Ideally, a comparison would be made of the perceptions of judges before and after the introduction of NCM, including performance-based budget systems in diverse judiciaries. However, it is not possible to retroactively gather data on the perceptions of judges before NCM. Furthermore, the implementation of NCM has been a gradual process that has taken place over many years. Therefore, we resort to comparing judiciaries with different, well-established levels of NCM, exemplified by their budgeting models. A general disadvantage of this method is that other differences, unrelated to NCM, may play a role. However, the influence of such differences is not obvious in this field. In principle, management, in particular NCM, can take case load (volume, complexity and size of cases) and the efficiency of procedural law into account. As a result, pressures on judges and their work related stress need not increase with
higher demands on the judiciary. We will address this issue in section 6. An advantage of this method is that judiciaries are compared at approximately the same moment in time. In view of these considerations, the main question that we endeavour to answer is how different management practices exert pressure on judges and impact their work-related stress.

3.1. Choice of countries

The judiciaries of three European countries, Italy, Finland and the Netherlands, have been selected following a diverse case method (Seawright and Gerring 2008) as they provide three examples of different types of court budgeting and management models, presumably leading to different behaviours of management. In Italy, the autonomy of the single judge extends to the organisation of his or her own work. The court president and the presidents of divisions were, until 2000, considered as “primus inter pares” with coordination functions, while rules and practices limited their powers. Between 2000 and 2016, there has been a general “change in the paradigm” (Vecchi 2018), driven by the Judicial Council, resulting in a different approach to performance measurement and management. In particular, court presidents were asked to develop strategies and plans to improve court performance and to establish targets at division and judge level. However, NPM principles are only now starting to make their way into the courts.

In Finland, before 2010, it was felt that courts were affected by problems of delay connected, among other factors, to the lack of cooperation of different actors (management, judges, lawyers etc.) and impractical resource allocation (Pekkanen 2011). In 2010, there was a large reform that reduced the number of courts and aimed at improving court organisation by creating larger units where work and resources could be organised more effectively, and where working methods and practices could be harmonised (Contini 2017). At the national level, a “management for results” approach has been implemented, including the setting of performance targets.

In the Netherlands, before 2000, the situation was similar to the other two countries: “judges generally worked as independent professionals, having much autonomy and leeway in determining their judicial work and not paying much attention to matters of management, organisation, money, or production” (Visser et al. 2019, p. 42). In 2002, three main changes were introduced in the courts: the creation of an executive board in charge of court management, the division of roles between judges and administrative staff which was put in charge of many organisational tasks, and the transition from input-based to output-based budgeting, while at the national level a Council of the Judiciary with broad responsibilities, including finance, was established.

Even if the development paths show similarities in the three countries (from judges’ individualism to more standardised and comprehensive managerial practices), the implications of the reforms are different, and the differences lie, in particular, in budgeting models and performance management. These budgeting models display three very different levels of the strength of the link between performance and funding. In Italy, there is no formal link between performance and funding, reflecting a very limited implementation of NCM. In Finland, a loose link (performance information is used as a basis for budgetary discussion, but it is not the only criterion taken into consideration). In the Netherlands, there is a close link. In Italy, programme budgeting
with no link between performance and budget allocation is in use. In Finland, performance targets are discussed by the Ministry of Justice and the courts and are negotiated together with the resources allocated by the Ministry of Justice to the courts. To measure the workload, and better assess, the budget needs a weighted caseload system is in use. The Netherlands bases the allocation of resources to courts on the formula $P \times Q$, where $P$ is the cost price of a range of categories of cases and $Q$ is the forecast of the quantity of resolved cases in these categories of cases.

As regards the setting of performance targets, in Finland national performance targets are set and approved together with the budget, while targets for individual courts are negotiated between the court and the Ministry of Justice (National Court Administration from 2021). The Ministry of Justice negotiates the resources to be allocated and the targets that can be achieved with these resources with each court. In Italy, there are no national performance targets, but divisional and individual targets are set by the Court President by taking into account the average productivity (number of adjudicated cases) of the previous three years. The production target for individual judges is confined to a range between -15% and +15% of the average, meaning that judges should not adjudicate less than 15% of the average but they should also not resolve more than 15% of the average. This “maximum requirement” was specifically requested by the judges to avoid too much pressure towards productivity.

In the Netherlands, the number of cases each court should adjudicate in a year is negotiated by the courts and the Council for the Judiciary, and it is directly related to the resources they will receive. The Council makes a forecast of the budget of the judiciary as a whole and the Ministry of Justice is bound by law to honour this proposal unless the finances of the State do not allow this. Courts that resolve more cases than forecasted, will receive 70% of the agreed price on the surplus of cases. Courts that produce less than expected must return 70% of the agreed price of the unsolved cases into an “equalisation account” managed by the Council for the Judiciary. Both in Finland and in the Netherlands, there are no formal individual targets, but all the judges are expected to help the court reach the overall targets. The characteristics of the three judiciaries are summarised in Table 1.

<table>
<thead>
<tr>
<th>Geographic location</th>
<th>Legal tradition</th>
<th>Link between performance and funding</th>
<th>Performance targets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Italy</td>
<td>Southern Europe</td>
<td>French</td>
<td>No link</td>
</tr>
<tr>
<td>Finland</td>
<td>Northern Europe</td>
<td>Scandinavian</td>
<td>Loose link</td>
</tr>
<tr>
<td>Netherlands</td>
<td>North Western Europe</td>
<td>French</td>
<td>Close link</td>
</tr>
</tbody>
</table>

Table 1. Main characteristics of the judiciaries participating in the survey.
3.2. Definition of “pressure on judges”

The caseload of judges has some resemblance with a gas: in the same way as a gas occupies all the space available, the caseload occupies all the available time of a judge.³ If the caseload is large, judges work under pressure to adjudicate as many as possible of the incoming cases. If the caseload is small, judges dedicate more time to study the cases and motivate the decisions. They still work under pressure, but it is possibly less intense. The implication is that spare time will not occur: judges are always busy working on their cases. It is not a 9 to 5 job, and working and studying after office hours is normal.

Against this background, a practical inspiration for the survey is the well-known Job Demands – Resources model (JD-R), used to assess work-related stress (Demerouti et al. 2001). Job demands refer to psychological, social and organisational aspects of the job that require psychological effort and are associated with costs (e.g. work pressure, work-related stress). At the opposite, job resources (job autonomy, social support) can mitigate work-related stress (Bakker and Demerouti 2007). This model has inspired a large survey that was conducted in 2013 in the Netherlands to assess the nature and development of judicial work among Dutch judges, after the “managerialisation” of the courts and, in particular, the transition from input-based budgeting to output-based budgeting (Fruytier et al. 2013, Visser et al. 2019). The results of that survey showed a general dissatisfaction with working conditions. In particular, judges complained about “increasing workload and decreasing quality, which were significantly related to job demands such as time pressure and production pressure and not sufficiently buffered by job resources”, together with a “creeping violation of their constitutional independence” (Visser et al. 2019, p. 49). While very relevant, this study misses a counterfactual, showing how an input-based budgeting system would have functioned under the then-prevailing circumstances or a comparison with other judiciaries. As noted in Section 2, dissatisfaction with working conditions and work intensity is common. In the present survey, we examine, therefore, work-related stress and pressure perceived by judges as a result of managerialisation by comparing three countries with different approaches to funding and managing the judiciary.

3.3. Variables

The main determinants of work-related stress examined here are:

- Physical pressure, in the sense of workload;
- Psychological pressure, exerted by management, colleagues and others;
- Stress reducers, such as motivation, autonomy and support.

Physical pressure consists of elements that increase the amount of time that judges need to dedicate to work. These elements are caseload, number of hearings, performance targets and timeframes. External factors play a major role (number of incoming cases), but other elements are often determined by court management (number of hearings, timeframes) or at the national level (court performance targets, linked to budgets). Psychological pressure is exerted by court management that may more or less strictly enforce performance targets, from other judges (peer pressure), from the Council of the

³ Based on conversations with Italian and Dutch judges.
judiciary or the Ministry of Justice, from parties and lawyers, and from public opinion. Both types of pressure directly affect stress, which also depends on the quality of working conditions (office space, ICT-tools and support staff) and tensions stemming from difficulties combining work and private life.

All elements contribute to work-related stress. Other elements can help judges to cope with stress: intrinsic motivation, professional standards that help to set priorities, organisational autonomy, and other stress reducers such as dialogue with and support from management, colleagues and administrative staff. These concepts can be represented in a simple model (see Figure 1).

3.3.1. Decomposition of the research question

To answer the main question formulated above, the following specific questions will be addressed.

1. Are judges working under pressure in the three judiciaries? Does this pressure affect the quality of their work?
2. Where is the focus of the management of the courts? Is it providing incentives to judges and other employees to reach performance targets? Does this reduce the professional autonomy of judges in their perception? And/or is the focus of management on organising the court in such a way that targets can be reached, thereby potentially decreasing the organisational autonomy of judges?
3. What is the impact of management on the workload of judges and on the pressures and stress that they experience?
4. Are professional standards of judges discernible and how do these standards, if any, interact with performance targets and efficiency in general? Do judges see performance targets set by management as binding?
5. Is the extrinsic and intrinsic motivation of judges related to the stress they experience, and what is the role of management in regulating stress?
3.3.2. Design and implementation of the survey

The survey was designed to answer these questions. Appendix 1 gives an overview of the content-related questions, clustered in broad categories. The survey was addressed to all first-instance judges working in courts of general jurisdiction in the three countries. In Finland, the questionnaire was sent by the National Court Administration to all judges of first-instance ordinary courts (about 500 judges). In the Netherlands, it was brought to the attention of all first-instance judges (about 2000 judges) by the Council for the Judiciary with the endorsement of the presidents of the courts and the explicit support of the National Association of Judges. In Italy, we sent the survey by personal e-mail to all presidents of the first-instance courts to inform them about the survey and, after a week, to all judges of these courts (about 4500 judges). As regards data protection, the data cannot be connected and traced back to the respondents in any way. Anonymity is guaranteed. By proceeding with the questionnaire, judges declared that they had been informed about the objectives and method of the survey. Table 2 provides key data about the survey.

### TABLE 2

<table>
<thead>
<tr>
<th>Country</th>
<th>Language</th>
<th>Time period</th>
<th>Total number of judges who received the questionnaire (approx.)</th>
<th>Number of respondents</th>
<th>Response rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Finland</td>
<td>English</td>
<td>29 June 2020 – 15 September 2020</td>
<td>500</td>
<td>73</td>
<td>15%</td>
</tr>
<tr>
<td>Italy</td>
<td>Italian</td>
<td>25 June 2020 – 15 September 2020</td>
<td>4.500</td>
<td>450</td>
<td>10%</td>
</tr>
<tr>
<td>Netherlands</td>
<td>English</td>
<td>23 November 2020 – 31 December 2020</td>
<td>2.000</td>
<td>198</td>
<td>10%</td>
</tr>
</tbody>
</table>

Table 2. Key data on the survey in three countries.

In the Netherlands the formal approval of the survey, in this case by the Council for the Judiciary, required more time than in the other countries. The questionnaire was submitted through Google Forms. It contained multiple choice questions (mostly on a Likert-scale), check box questions, short answer questions and open questions (short paragraphs).

3.3.3. Response rate

Table 2 gives the response in absolute and relative terms. The response rates in the three countries are sufficient for outcomes to be representative, assuming that selection effects play a limited role. For the comparison of the three judiciaries which is the focus of this article, it is sufficient that, if any, selection effects are similar. However, there are no indications that selection effects occur. Section 6 shows that the views on work-related

---

4 The complete survey is available on [www.uu.nl](http://www.uu.nl)
5 The ENCJ conducts a survey among the judges of Europe every three years on the independence of the judiciary. The last survey was held in the first quarter of 2022 with response rates of 19% in Finland, 12% in
stress are homogeneous among the respondents, implying that if selection effects occur with respect to the personal characteristics of judges, these have a small impact.

### 3.3.4. Timing of the survey in relation to the COVID Pandemic

The survey took place after the closure of the courts due to the COVID Pandemic. In the Netherlands, the courts physically closed in March 2020 and started conducting oral hearings again in April 2020, where needed. In the following months oral hearings increased, and the situation evolved towards normalcy. In Finland, physical presence at the courts was limited and many hearings were suspended in March and April 2020. In Italy, from March 2020 to May 2020 courts were closed and all hearings were postponed. From May to June, access to courts was limited and hearings were conducted, when possible, remotely.

We conclude that in all three countries, normal operations at the courts had resumed in June in a similar way. In the survey, respondents were asked to consider the last two years when answering the questions. The pandemic may still have caused more noise in the data than normally expected. In presenting the results, we therefore focus on major differences among the three judiciaries.

### 3.4. Statistical analysis

In the quantitative analysis, data were analysed using Excel, Tableau and SPSS. First, a descriptive analysis was performed to determine average scores, standard deviations and frequencies. Secondly, per question, we used single-factor (one-way) ANOVA to test the null hypothesis that the means of the three countries are the same. When these F-tests (or in the case when homogeneity of variances is not met, Welch-test) show that the means of the three countries are significantly different, we performed Tukey’s honestly significant difference (Tukey HSD) post hoc test (or in the case when homogeneity of variances is not met, Games-Howell) to determine which of the countries differ. Section 6 uses ordinal regression.

The open questions have been treated as qualitative data. They have been rearranged and synthesised to aggregate similar comments that support the same view. The most explicative comments among those reflecting the same opinion were selected, and these are cited in italics in the next section to support the quantitative analysis.

### 4. Results

#### 4.1. General outcomes

The answers show differences as well as similarities among the three judiciaries. The results of the survey exhibit large differences among the three judiciaries in terms of work organisation, setting of individual performance targets, judges’ autonomy, dialogue between managers and professionals, and focus on quality. The answers show similarities, especially regarding work-related stress, perceived pressure and motivation. In this section we will first examine the characteristics of the respondents.
We then address workload, as a difficult to observe but important variable, against which background the views of judges should be interpreted. After that, we address the questions presented in the previous section.

4.2. Characteristics of respondents

Figure 2 provides an overview of the characteristics of the respondents. Panel 1 concerns the area of law they work in. Differences in the legal system come to light in this panel. In the Netherlands, administrative law is a regular part of the work of the first-instance courts, while in the other countries these cases are handled by separate courts. In Finland many judges are not specialised and handle both criminal and civil cases. In Italy relatively many presidents of (sections of) courts were among the respondents. Panel 2 reflects on the one hand, that many judges in the Netherlands work part-time,\(^6\) and on the other hand that many judges, including part-time judges, work overtime. Panels 3 and 4 show that judges in the Netherlands have more experience and are older than in Italy, followed by Finland. The gender distribution, not depicted in Figure 2, is similar (49.6% women in Italy, 51.5% in Finland and 56.4% in the Netherlands).

---

\(^6\) In the Netherlands, professional judges can work full time or part time, with no impact on their terms of employment. In 2022, the total number of judges was 2,636, in full time equivalents 2,451 (Raad voor de rechtspraak 2023).
FIGURE 2

1: field of work

![Field of Work Diagram]

- Administrative judge
- Criminal Judge
- Judge in criminal and non-criminal cases
- Non-criminal Judge
- President of the court or section/division
- Other

2: hours worked

![Hours Worked Diagram]

- Less than 30 years old
- 30–40 years old
- 40–50 years old
- 50–60 years old
- More than 60 years old
3: years of experience

![Years of Experience Graph]

4: age

![Age Distribution Graph]

Table 2. Key data on the survey in three countries.

4.3. Question 1: are judges working under pressure?

To get an impression of the workload of judges in the three countries, we first examine the volume of the cases they handle in a year. Comparing the caseload of judges in the three countries is a difficult task, and the risk is that any comparison is misleading. The weighting systems to account for the large variety of court cases are different in the three countries, and cannot be used to compare caseload.
FIGURE 3

1: I am always well prepared for hearings

![Bar chart showing the percentage of respondents who strongly agree, agree, neutral, disagree, and strongly disagree with being well prepared for hearings.](chart1.png)

2: I often feel forced to excessively simplify the reasoning of my written judgments, in order to dispose enough proceedings

![Bar chart showing the percentage of respondents who strongly agree, agree, neutral, disagree, and strongly disagree with the necessity to simplify reasoning.](chart2.png)

Consequently, we have to make do with unweighted figures to get an indication of differences between countries. To get reliable results, the case distribution within areas of law must be roughly the same. In Finland, where a judge often deals with both criminal and civil cases, the median per year of adjudicated criminal and civil cases is 100 cases per judge, while the median of pending cases is 60 cases per judge. As regards non-criminal cases (excluding administrative cases), in Italy and in the Netherlands the median for the number of adjudicated cases is similar: 250 resolved cases per judge per
year in Italy and 200 in the Netherlands,\(^7\) while the number of pending cases is very different: in Italy the median is 550 pending cases per judge, in the Netherlands it is 20. As regards criminal cases, in the Netherlands there are 360 resolved cases per judge and 10 pending cases, in Italy 250 resolved cases and 300 pending cases. Roughly speaking, judges in Italy and the Netherlands handle more cases per year than those in Finland. In Italy, work is burdened by enormous backlogs of cases that require the occasional attention of judges.

It should be noted that workload and caseload are not the same. Besides hearing and deciding cases, judges perform different tasks in the three countries: in the Netherlands many tasks are dealt with by court staff, while in the other two countries, judges for example schedule their hearings themselves. In the next section, we return to this issue. We tentatively conclude that workload is highest in Italy (further IT), followed by the Netherlands (NL) and then Finland (FI).

Whether a high workload is problematic can be gleaned from the consequences it has for the quality of the work. Several items of the survey shed light on the quality of the work judges believe they are delivering. When asked whether they have enough time to study cases properly, high percentages of judges answer that they do not: 33% in NL, 42% in IT and 51% in FI (differences between the means\(^8\) are not significant). A second item concerns whether the respondents are always well prepared for hearings: 84% of the respondents in NL, 59% in IT and 52% in FI believe they are. See Figure 3. The differences between the means of NL and IT, as well as NL and FI are significant (\(p=0.0000\)), while the difference between IT and FI is not significant. The third item addresses whether the respondent is often forced to excessively simplify the reasoning of his/her written judgments, in order to dispose of enough procedures: 16% in NL, 37% in IT and 40% in FI do so. The differences of the means between NL and IT, as well as NL and FI are significant at, at least, 5% level. The difference between IT and FI is not significant. These results indicate that, according to the judges, quality is least under pressure in NL, while there are no significant differences between IT and FI.

Among the three countries, in IT the workload is considered by the respondents least sustainable, although the differences are not large: 38% of the judges declared workload to be unsustainable versus 32% in NL and FI. The average score for sustainability is 2.8 in IT on a scale from 1 (absolutely unsustainable) to 5 (absolutely sustainable), while in the NL the average score is 3.1 and in FI 3.0.\(^9\) The difference between the NL and IT is significant at 1% level. Furthermore, in IT 87% of respondents consider the caseload as a source of stress (84% in the NL, 75% in FI), where the difference of the means for NL and IT is significant at 1% level, and for FI and IT at just above 5% level (\(p=0.052\)).

These outcomes would lead ceteris paribus to work-related stress to be lowest in NL. This is partly found. On a five-point scale between 1 (low stress) and 5 (high stress), respondents rate their work-related stress at 3.4 in NL and FI and 3.8 in IT. Differences

\(^7\) As noted above, in the Netherlands many judges work parttime. As a result, the reported number of cases underestimates the case load of fulltime judges.

\(^8\) Comparison of the means per country of the weighted percentages of generally five answer categories.

\(^9\) As abilities differ among judges, whether a judge considers a workload is sustainable depends on his/her abilities if cases are apportioned uniformly. If nearly 40% as in Italy considers the caseload unsustainable, this points to a real issue.
of the means of NL and FI compared with IT are significant at 1% level. The difference between NL and FI is not significant. It should be emphasised that in all three judiciaries stress is quite high. For instance, the maximum level of stress is reported by 11% of the respondents in NL, 24% in IT and 10% in FI.

We conclude that workload is high in all three judiciaries and foremost in Italy. This puts pressure on judges and leads to quality issues. In this situation, it is of interest to examine what court management is focusing on.

\textbf{FIGURE 4}

\textit{1: There are standard instructions/practices/rules to schedule hearings}

\textit{2: I am autonomous in scheduling my hearings}

\textbf{Figure 4. Organisational autonomy.}
4.4. Question 2: where is the focus of management?

The second question we posed in the introduction of this article is the role and focus of management in the three judiciaries. The main potential difference in focus is between the organisation of the courts and its work processes and procedures versus the improvement of the productivity of personnel, given the organisation of the court. The latter is about performance management and individual incentives. Both approaches are inspired by NPM. The three judiciaries differ much in this respect.

4.4.1. Organisational perspective

In the Netherlands, managerial control is evident when it comes to work processes. For instance, 90% of the respondents agree with the statement “There are standard instructions/practices/rules to schedule hearings” (FI 59%, IT 70%). Consistent with this, only 18% declare to be autonomous in scheduling their hearings (89% in IT and 92% in FI). See Figure 4. Among the judges of the three countries, Dutch judges are by far the least autonomous in this respect. On a scale from 1 to 5 (where 5 is the maximum level), the self-estimated autonomy in organising one’s work obtained an average score of 3.5 in the Netherlands, 4.2 in Italy and 4.4 in Finland. Only 36% of Dutch respondents declared to be autonomous in deciding which cases to deal with first (vs 80% in IT and 86% in FI), while 62% declared to be autonomous in deciding how to organise their work (90% in IT and 93% in FI).

The reasons for this striking difference lie in the different internal organisation of Dutch courts, especially as regards the role of support (non-judge) staff in dealing with organisational tasks. Support staff, in fact, is in charge of scheduling hearings for the cases on the basis of a standardised calendar (days of hearings and the number of cases per hearing are fixed). Therefore, judges do not decide which cases to deal with first. On the one hand, this reduces the organisational autonomy of judges, on the other hand, judges can delegate all administrative tasks, having more time to spend on hearing and deciding cases. This contrast is well explained in the following comment, as it was expressed literally:

I sometimes feel like I work on an assembly line where hearings are just appearing in front of me the entire time, which can feel like a lot of pressure. On the other hand, I feel ‘blessed’ (compared to judges in other countries) that I don’t have to spend time scheduling my cases / hearings: a lot of organisational work is taken out of my hands, so I have more time to spend on the cases itself. There are weeks where the ‘assembly line’ is causing a problem: when a casefile or hearing is costing more time than is scheduled, and other hearings are already appearing. This then leads to working free [i.e. spare] time.

In the other two countries, many respondents would like to have more common practices among judges (71% of respondents in IT, 64% in FI, in contrast to 38% in NL). A Finnish judge adds this comment:

I think we could be less autonomous and have more standard practices that each judge would be required to follow. Not to in anyway influence the rulings but just the way how the case is processed forward and, ideally, length of proceeding. That would give

---

10 All differences of the means of NL on the one hand and FI and IT on the other hand are highly significant.
better predictability on the timing etc. to parties and perhaps help advocates to plan their work and lower the costs of litigation.

In Italy, several judges complain about disorganisation and excessive workload:

The problem is not autonomy in organisation, but poor general organisation that forces almost everyone to work in emergency conditions. Moreover, in the organisation, the constraints imposed by the calendar of hearings fixed for the whole court on the basis of the available courtrooms and the needs of administrative staff cannot be overcome. Organisational autonomy is strongly conditioned by the shortage of available courtrooms and administrative staff.

4.4.2. Individual performance perspective

The three countries differ in the use of individualised production targets. In IT, 66% of the respondents answer that individual performance targets are clear and specific, while in FI only 22% and in the NL 35% think so (Figure 4). In IT 67% of the respondents agree with the statement that performance targets of courts and court divisions are well known by all judges, 70% agree that these targets are accepted by most judges, and 51% sees them as formal arrangements.
In the Netherlands, instead, there are no formal individual performance targets (only 28% perceives targets as formal). “Targets are set by management for the entire department. In addition, there are professional standards (agreed upon in expert meetings [of judges]) which define caseload for judges. These are applied for scheduling hearings of each judge.” Furthermore, “each judge has to do a certain amount of hearings every week” (during 2 days per week) and “there are performance targets in the number of (weekly) hearings; not in case numbers”.

Figure 5. Individual performance targets.
In Finland, targets are set at section level. “Ministry and nowadays National Court Administration make Performance targets agreement with chief judge of my court and he makes targets to every section of court. Officially there is no individual target of individual judge but of course that system needs that everybody makes his shares to sections target.”

In Italy, courts and division targets are set by court presidents, taking the maximum workload required of the judges into account, and in discussion with them. “Performance targets are indicative and formulated on the basis of the previous year’s average”: they are calculated on the basis of the average number of adjudicated cases in the past three years, and they must remain in a range between -15% of the average to +15%, as noted earlier.

Apparently, Italy is the only country where individual performance targets are clear and specific. In the other two countries performance targets are considered as informal, and these are more like a moral obligation. In addition to production targets, there are timeliness standards. We will discuss these in the next section.

4.4.3. Summing up

There is a different focus of management in the three countries. In NL court management is focused on the efficient organisation of work. Due to the division of labour, the production process resembles, one might say, as the respondent quoted above does, an assembly line. Performance targets are largely implicit. In FI and IT judges organise their own work, either without management setting explicit production targets (FI) or with management doing this (IT). The question arises whether explicit or implicit performance management exerts more pressure on judges.

4.5. Question 3: does performance management exert pressure on judges?

We saw that performance management plays a role, but in a very different manner. To examine the impact of performance management, a first question is whether performance targets are difficult to meet for most judges. If not, pressure is low and impacts only dysfunctional judges. Performance targets are considered “too ambitious” by 27% of Italian respondents, 34% of Finnish and 53% of Dutch respondents (differences of means of NL and IT as well as NL and FI are highly significant). It seems that in Italy and Finland’s targets are relatively easy to reach. As to timeliness, 66% of the respondents in Italy, 57% in NL and 54% in Finland agree that the targets for the timeframes of judicial procedures are difficult to meet. Timeliness seems to be more of a challenge than production as such.

Having (ambitious) targets is one thing, enforcing them is another. Turning to enforcement, 63% of the respondents in IT agree that section and court presidents check that all judges meet performance standards against 40% in FI and 27% in NL (means differ significantly between NL and IT and NL and FI at 1% level; no difference between IT and FI). A judge’s failure to meet targets has significant consequences for the judge, according to 31% of the respondents in IT, 22% in FI and 11% in NL. Only the difference between NL and IT is significant (p=0.0000).

Similarly, the court/section president may take action if the length of proceedings is about to exceed the target. In NL 26% perceive presidents to actually take action, 35% in
Italy and 42% in Finland (mean significant between IT and NL and FI and NL at 5%). There are significant consequences for judges if the targets for the length of proceedings are not met, according to only 3% of the respondents in NL, 29% in Italy and 16% in Finland (all differences significant at least at 5% level). Table 3 summarises.

### TABLE 3

<table>
<thead>
<tr>
<th></th>
<th>FI</th>
<th>IT</th>
<th>NL</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Production targets are too ambitious</strong></td>
<td>34%</td>
<td>27%</td>
<td>53%</td>
</tr>
<tr>
<td><em>(Sign. difference: IT-NL, FI-NL)</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>President is monitoring whether a judge meets his/her production targets</strong></td>
<td>40%</td>
<td>63%</td>
<td>27%</td>
</tr>
<tr>
<td><em>(sign. difference: IT-NL, FI-NL)</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Sign. consequences for a judge if s(he) does not meet production targets</strong></td>
<td>22%</td>
<td>31%</td>
<td>11%</td>
</tr>
<tr>
<td><em>(sign. difference: IT-NL)</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Timeliness standards are difficult to meet</strong></td>
<td>54%</td>
<td>66%</td>
<td>57%</td>
</tr>
<tr>
<td><em>(no sign. differences)</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>President takes action if timeliness targets are about to be exceeded</strong></td>
<td>42%</td>
<td>35%</td>
<td>26%</td>
</tr>
<tr>
<td><em>(sign. difference: IT-NL, FI-NL)</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Sign. consequences for a judge if s(he) does not meet timeliness targets</strong></td>
<td>16%</td>
<td>29%</td>
<td>3%</td>
</tr>
<tr>
<td><em>(significant difference: IT-FI, IT-NL, FI-NL)</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Control/pressure by president affects my independence in a negative way</strong></td>
<td>25%</td>
<td>13%</td>
<td>22%</td>
</tr>
<tr>
<td><em>(no sign. differences)</em></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3. Respondents who agree with statements about performance targets, as percentage of respondents who expressed an opinion on these statements. (Note: significant difference p < 5%).

We find that, according to the respondents, production targets are not overly ambitious, especially in Italy and Finland. Monitoring of the individual performance of judges by the president is not common, and sanctioning of judges is even less so. The same holds with regard to timeliness. Thus, explicit performance management is not very aggressive. While performance management could conflict heavily with judicial independence, in practice this is not much of issue in all three judiciaries. See Table 3, last row, where the differences between countries are insignificant. It should be stressed that pressure on judges may not only stem from monitoring and sanctioning by management, but also by management making a moral appeal on the responsibility of judges for the functioning of the court (NL). It should be noted as well that production targets may also serve to protect judges from overload or help them deal with the overload. In Italy production targets also set a maximum which can be explained by the enormous backlog of cases (see above). It would be better for judges to drastically reduce the backlog, but if that is not possible due to lack of resources, the only thing

---

11 This excludes respondents who answered ‘not available’ or ‘not applicable’. The number of respondents who express an opinion varies across the statements: 56-65 for FI, 419-432 for IT and 166-185 for NL. In particular, timeliness standards are not recognized as such by a substantial number of respondents in NL and FI.
management can do is to protect judges from overload by setting an upper bound for production.

4.6. Question 4: are professional standards discernible and do they interact with performance targets?

An insight into the professional standards of judges can be derived from the priorities judges set for themselves. Their opinion was asked about the statement: “I give the needed attention to each case, even if this may decrease the court’s efficiency”. In NL 89% agree, 88% in FI and only 50% in IT (Figure 6). Similarly, on the statement “I give the needed attention to each case, even if this may increase the time of disposition” 92% in NL, 89% in Finland and 65% in Italy agreed. Differences of the means are significant between IT and NL and IT and FI (p=0.0000). In line with these priorities, only 21% of the respondents in NL, 20% in FI but 53% in IT perceive the court’s targets as binding (differences between IT and the other two countries are significant (p=0.0000)). Legal quality is an essential aspect of the professional standards. As noted before, only a minority of respondents feel forced to excessively simplify the reasoning of their written judgments in order to dispose of enough proceedings: 16% in NL, 37% in Italy and 40% in Finland. An indication of the behaviour towards each other is the opinion of the respondents that most judges do not monitor each other’s behaviour when it comes to the fulfilment of performance targets: 86% in NL, 49% in IT and 64% in FI (means differ significantly at 5% level). While the response of Italian judges is likely to be driven by necessity (backlogs, lack of funding), efficiency and timeliness do not override typical judicial values such as legal quality and integrality. Given the relatively weak enforcement of performance standards that we noted above, there does not seem to be a strong clash between the professional standards of the judges and the values of court management. This may not be surprising as management consists of judges.
FIGURE 6

1: I give the needed attention to each case, even if this may decrease the court’s efficiency

2: I give the needed attention to each case, even if this may increase the time of disposition

4.7. Question 5: is extrinsic and intrinsic motivation related to stress and what is the role of management in regulating stress?

Tables 4 and 5 present a broad perspective on factors that lead to work-related stress and factors that reduce stress. As it is the task of management to ensure that the “job” gets done in a sustainable way, the influence of management on the stress that judges (and
other personnel) experience is of particular interest. The tables show that judges in IT work in nearly every aspect under higher pressure than judges in the other countries. For all the judges in the three countries, the most stress-generating element is, without a doubt, caseload, followed by timeliness targets. This result confirms what was found by the previous studies (see Casaleiro 2021).

Apart from these factors, the contribution of management and other governing bodies as such to stress is rather limited. The parties and their lawyers put substantial pressure on judges in all three judiciaries, while public opinion is felt in IT and NL in particular. Another source of stress in IT is the lack of facilities for judges (office space, IT, support staff). We saw earlier that especially in NL the focus of management is on facilitating judges. In the other judiciaries (in particular IT) management takes this role less actively or effectively. Still, the best management can do is to keep the caseload workable for the judges by maintaining an adequate balance between production capacity (number of judges and support staff) and caseload.

**TABLE 4**

<table>
<thead>
<tr>
<th>Work-related stress is caused by:</th>
<th>FI</th>
<th>IT</th>
<th>NL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Caseload (sign. difference IT-NL)</td>
<td>75%</td>
<td>86%</td>
<td>83%</td>
</tr>
<tr>
<td>Performance targets (sign. difference IT-FI, IT-NL)</td>
<td>32%</td>
<td>55%</td>
<td>32%</td>
</tr>
<tr>
<td>Timeframes (length of proceedings) targets (sign. difference IT-NL)</td>
<td>55%</td>
<td>69%</td>
<td>50%</td>
</tr>
<tr>
<td>Pressure from the president (section or court) on the organisation of work (no sign. difference)</td>
<td>14%</td>
<td>19%</td>
<td>14%</td>
</tr>
<tr>
<td>Pressure from colleagues (sign. difference FI-NL, IT-NL)</td>
<td>5%</td>
<td>7%</td>
<td>10%</td>
</tr>
<tr>
<td>Pressure from public opinion (sign. difference IT-FI, FI-NL)</td>
<td>25%</td>
<td>40%</td>
<td>41%</td>
</tr>
<tr>
<td>Pressure from the Judicial Council / National Court Administration (sign. difference IT-FI, FI-NL)</td>
<td>10%</td>
<td>22%</td>
<td>28%</td>
</tr>
<tr>
<td>Pressure from the Ministry of Justice (no sign. differences)</td>
<td>11%</td>
<td>14%</td>
<td>18%</td>
</tr>
<tr>
<td>Pressure from parties and/or lawyers (no sign. differences)</td>
<td>47%</td>
<td>55%</td>
<td>46%</td>
</tr>
<tr>
<td>Lack of adequate working space (sign. difference IT-NL, FI-NL)</td>
<td>21%</td>
<td>64%</td>
<td>25%</td>
</tr>
<tr>
<td>Lack of adequate ICT tools (hardware, software) (sign. difference IT-NL)</td>
<td>55%</td>
<td>51%</td>
<td>42%</td>
</tr>
<tr>
<td>Lack of adequate administrative support (clerks, judicial assistants etc) (sign. difference IT-NL)</td>
<td>52%</td>
<td>62%</td>
<td>53%</td>
</tr>
<tr>
<td>Difficulty of work-private life balance (sign. difference IT-FI, IT-NL)</td>
<td>44%</td>
<td>59%</td>
<td>49%</td>
</tr>
</tbody>
</table>

Table 4. Causes of work related stress, respondents that experience pressure leading to stress as percentage of all respondents.
(Note: significant difference p < 0.05. N is 73 for FI, 448 for IT and 195 for NL.)
TABLE 5

<table>
<thead>
<tr>
<th>Factors that help to relieve work-related stress</th>
<th>FI</th>
<th>IT</th>
<th>NL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Support from administrative staff (sign. difference IT-FI, FI-NL)</td>
<td>73%</td>
<td>90%</td>
<td>93%</td>
</tr>
<tr>
<td>Support from colleagues (sign. difference IT-FI, IT-NL)</td>
<td>92%</td>
<td>79%</td>
<td>93%</td>
</tr>
<tr>
<td>Support from the president (sign. difference IT-FI, IT-NL)</td>
<td>66%</td>
<td>82%</td>
<td>47%</td>
</tr>
<tr>
<td>Strong motivation (sign. difference IT-FI)</td>
<td>99%</td>
<td>92%</td>
<td>96%</td>
</tr>
<tr>
<td>Perception of my role as a judge in the community (no sign. differences)</td>
<td>78%</td>
<td>84%</td>
<td>87%</td>
</tr>
<tr>
<td>Autonomy / freedom in managing my time and priorities (no sign. differences)</td>
<td>96%</td>
<td>92%</td>
<td>94%</td>
</tr>
<tr>
<td>Level of salary (sign. difference IT-FI, IT-NL)</td>
<td>49%</td>
<td>67%</td>
<td>49%</td>
</tr>
</tbody>
</table>

Table 5. Mitigating factors of work-related stress, respondents that experience relief as percentage of all respondents.
(Note: significant difference p < 0.05. N is 73 for FI, 448 for IT and 195 for NL.)

Examining the factors that relieve stress, the element that contributes most to the reduction of work-related stress is strong motivation, together with autonomy in managing time and priorities, while the element which contributes least to work-stress relief is the level of salary. In Italy, an important role is played by the President of the court and his/her support of the judges, while in the other two countries, especially in the Netherlands, this support is perceived as much less important than other factors.

While it is not obvious by which mechanism strong motivation and one’s role in society leads to stress reduction, this is, in any case, what keeps the judges going, in combination with their autonomy and freedom in managing time and priorities. Essentially, this reflects a very strong intrinsic motivation of the judges in all three judiciaries, and all three rely on this intrinsic motivation. Extrinsic incentives are, according to the results that we presented, relatively weak. Nevertheless, performance targets do cause stress.

5. Differences among judges

For the interpretation of the results, it is relevant to examine whether judges form a homogeneous group or display significant differences. We limit ourselves to the essentials, i.e. work-related stress. Table 6 presents work-related stress by characteristics of the respondents. For ease of presentation, significance is calculated relative to all observations per country and in total. Apart from hours worked, which is not a fixed characteristic, there are few differences among judges. The differences concern gender, with female judges experiencing significantly more work-related stress than male judges in Italy and overall. In the other two countries, there is also a difference, but it is not significant.

As to age, there are a few significant differences. Judges older than 60 have significantly less stress than all judges together in the Netherlands and overall, while the few judges younger than 30 experience more stress. Area of law and experience as a judge do not show differences in stress levels. Where differences are significant, these are not particularly large.
Hours worked captures contracted hours (including part-time judges) and overtime. It is therefore not a fixed characteristic, shows much more variation in Italy and Finland. The differences in the Netherlands are surprisingly small, except for judges that work more than 50 hours per week.

### Table 6

<table>
<thead>
<tr>
<th></th>
<th>Italy</th>
<th>Finland</th>
<th>Netherlands</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td><strong>Gender</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Men</td>
<td>3.6*</td>
<td>1.0</td>
<td>3.2</td>
<td>1.1</td>
</tr>
<tr>
<td>Women</td>
<td>4.0*</td>
<td>0.8</td>
<td>3.6</td>
<td>0.9</td>
</tr>
<tr>
<td><strong>Age class</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 30 years</td>
<td>0.0</td>
<td>0.0</td>
<td>4.0*</td>
<td>0.0</td>
</tr>
<tr>
<td>30 - 40 years</td>
<td>3.8</td>
<td>0.9</td>
<td>3.5</td>
<td>1.1</td>
</tr>
<tr>
<td>40 - 50 years</td>
<td>3.8</td>
<td>0.9</td>
<td>3.4</td>
<td>1.0</td>
</tr>
<tr>
<td>50 - 60 years</td>
<td>3.8</td>
<td>0.9</td>
<td>3.3</td>
<td>1.1</td>
</tr>
<tr>
<td>&gt; 60 years</td>
<td>3.8</td>
<td>1.0</td>
<td>2.9</td>
<td>0.8</td>
</tr>
<tr>
<td><strong>Experience</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 5 years</td>
<td>3.9</td>
<td>0.9</td>
<td>3.6</td>
<td>1.0</td>
</tr>
<tr>
<td>5 - 10 years</td>
<td>3.9</td>
<td>0.8</td>
<td>3.8</td>
<td>0.8</td>
</tr>
<tr>
<td>&gt; 10 years</td>
<td>3.8</td>
<td>0.9</td>
<td>3.1</td>
<td>1.0</td>
</tr>
<tr>
<td><strong>Area of law</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-criminal</td>
<td>3.8</td>
<td>0.9</td>
<td>3.4</td>
<td>0.9</td>
</tr>
<tr>
<td>Criminal</td>
<td>3.8</td>
<td>0.9</td>
<td>3.3</td>
<td>1.1</td>
</tr>
<tr>
<td><strong>Hours/week</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 35</td>
<td>2.7*</td>
<td>1.3</td>
<td>1.3*</td>
<td>0.5</td>
</tr>
<tr>
<td>35 - 40</td>
<td>3.4*</td>
<td>0.9</td>
<td>2.7*</td>
<td>0.8</td>
</tr>
<tr>
<td>40 - 45</td>
<td>3.6*</td>
<td>0.9</td>
<td>3.5</td>
<td>0.7</td>
</tr>
<tr>
<td>45 - 50</td>
<td>3.9</td>
<td>0.8</td>
<td>3.9*</td>
<td>0.6</td>
</tr>
<tr>
<td>&gt; 50</td>
<td>4.3*</td>
<td>0.8</td>
<td>4.1</td>
<td>1.2</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>3.8</td>
<td>0.9</td>
<td>3.4</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Table 6. Mean work-related stress by personal characteristics.
(Note: p < 0.05% of difference between group and total per country or overall denoted by *.)

### 6. Determinants of work-related stress

While the emphasis is on the differences among countries, we have examined connections between the answers, applying the scheme of section 3 (Figure 1). The table in Appendix 2 gives the outcome of an ordinal regression analysis with work-related stress as the dependent variable. The results are confined to the variables that are significant for at least one of the ordinal classes. As to the characteristics of the respondents, only gender is significant. Not surprisingly, after the previous section, hours worked have an impact. Judges not having enough time to study cases properly has a large influence on stress. Also, the answer to the question of whether failure to meet targets has significant consequences for the judge is important for stress. Pressure from parties and/or lawyers and lack of administrative support has significant influence in some ordinal classes but has less impact. Finally, the two country dummies are
significant, implying that not all differences are covered by the other variables. This outcome warrants further research.

It is also of interest that other variables are not significant. For instance, organisational autonomy has no impact on stress. Thus, freedom to organise one’s work does not lead to less stress, or it is compensated by the effort it takes to organise one’s work. Also, the clarity and specificity of performance standards is not relevant for stress.

7. Conclusions

In the last decades, many judiciaries underwent the application of private sector techniques to courts of justice (NCM), including the appointment of managers within the courts, the re-organisation and optimisation of services, the use of performance measures and management and the implementation of performance-based budgeting models. These instruments interact with the professional standards of judges that emphasise independence and quality. The main question that we attempt to answer is how different management practices exert pressure on judges and impact their work-related stress. To address this question, a survey was conducted among the judges of ordinary, first instance courts in Finland, Italy and the Netherlands. In these three countries, different budgeting models are in place, reflecting different degrees of implementation of NCM-methods. The Netherlands has the highest degree and Italy the lowest. Finland takes a middle position. To summarize the results, we answer the five sub questions that were formulated.

1. Are judges working under pressure in the three judiciaries and does this pressure affect the quality of their work? The answer is yes. In all three judiciaries, especially in Italy, high workload puts a lot of pressure on judges and the quality of their work suffers as a result. The lowest impact on quality is found in the Netherlands.

2. Where is the focus of the management of the courts? There is a striking difference between the Netherlands and the other two judiciaries. In the Netherlands, the focus of management is on the internal organisation of the courts, especially as regards the role of support (non-judge) staff in dealing with organisational tasks. Support staff is, for instance, in charge of scheduling hearings. As a result, the organisational autonomy of judges is low. Performance targets are largely implicit. In Finland and Italy judges organise their own work, either with (Italy) or without (Finland) management setting explicit production targets. Italy is the only country where individual performance targets are clear and specific. In the other two countries, performance targets are considered informal and more like a moral obligation.

3. What is the impact of management on the workload of judges and on the pressures and stress that they experience? According to the respondents, production targets are not overly ambitious, especially in Italy and Finland. Monitoring of the individual performance of judges by the president is not common, and sanctioning of judges is even less so, least in the Netherlands. The same holds true for timeliness standards. Thus, explicit performance management is not aggressive at all. While performance management could conflict heavily with judicial independence, in practice this is not a serious issue in all three judiciaries.
4. Are professional standards of judges discernible and how do these standards, if any, interact with performance targets and efficiency in general? An insight into the professional standards of judges can be derived from the priorities judges set for themselves. Legal quality takes precedence over efficiency and timeliness, in particular in the Netherlands and Finland. This shows, for instance, in the vast majority of respondents giving the needed attention to each case, even if this decreases the court’s efficiency or increases the time to disposition. In Italy, only between 50% and 65% do so. In line with these priorities, only few respondents in the Netherlands and Finland perceive performance targets as binding, whereas a small majority in Italy do.

5. Is the extrinsic and intrinsic motivation of judges related to the stress they experience, and what is the role of management in regulating stress? The survey shows that in all three judiciaries the intrinsic motivation of judges is very strong and essentially keeps them going under high pressure. Extrinsic incentives are, according to the results that we presented, relatively weak. Nevertheless, performance targets do cause stress. In Italy, an important role is played by the president of the court and their support for the judges, while in the other two countries, especially in the Netherlands, this support is perceived as far less important than other factors.

In all three judiciaries, workload and work stress are high, regardless of the court organisation, level of managerial control and budget system, but the levels differ. In NCM a major role is played by performance management with respect to caseload and timeliness. However, in all three judiciaries most of the pressure on judges is not caused by performance management: production targets are not ambitious, or if they are relatively ambitious, the targets are not enforced, and there are no serious consequences for judges that fail to meet the standards. In Italy, targets even protect against overload caused by backlogs and timeliness issues. This implies that there are not many tensions between the demands for independence and production. The other side of the coin is that performance management is largely ineffective in dealing with backlogs and long disposition time. This outcome honours the independence of the judge, but it does not resolve the performance issues of the judiciary, where these exist.

Nonetheless, NCM can also focus on the optimal design of work processes and the division of labour between judges and administrative/legal staff. This reduces the organisational autonomy of judges but delivers them from administrative tasks and improves the predictability of procedures for the parties. The Netherlands provides an example where judges do not have individualised production targets, but their work is regulated by the number of hearings. They report that it feels like an assembly line with high implicit production targets. Still, the assembly line can be stopped by the judge if the need arises. Quality standards stipulate that when more time is required for a case, this need transcends efficiency and timeliness. Strict output budgeting fits into this organisational model, as it profits from the high predictability of procedures.

The development of professional (quality) standards by the judges themselves seems to be triggered by increasing management pressure to achieve efficiency. Professional standards are designed as an instrument for judges and not for management: they embody the judges’ vision of quality standards (see also Contini 2017), and management has to take these standards into account. These professional standards seem to be rooted in the Dutch judges’ culture by now, and this is evident from the answers provided in
the survey. What once were individual opinions, as they still largely are in the other two countries, have developed into professional standards, and this process seems to have led to a stronger position of judges vis a vis management.

Our study and, in particular, the comparison of Italy and the Netherlands but also of Finland and the Netherlands, indicate that, while not a magic solution, NCM does not lead to higher pressure and stress than the traditional management methods of the courts. Indeed, it may help to reduce pressure by improving the organisation of the courts and making professional standards and financial needs more explicit.

References


De Santis, L.G., Emery, Y., and Egloff, L., 2016. Culture (sub-project 5). In: A. Lienhard and D. Kettiger, eds., The judiciary between management and the rule of law, Results of the research project Basic research into court management in Switzerland. Bern: Stämpfli/Nomos/Jan Sramek.


## Appendix 1. Content-related survey questions, organised by topic

<table>
<thead>
<tr>
<th>Topics</th>
<th>Questions and statements</th>
</tr>
</thead>
</table>
| Physical pressure             | How many cases did you solve last year (approximately)?  
How many pending cases do you have as of today (approximately)?  
How many hours per week do you work on average?  
How many hours per week do you spend on hearings?  
Are there individual performance targets for you as a judge? Who did establish them?  
Is the length of proceedings monitored? By whom?  
In addition to the priorities set by the law, are there other standards/timeframes/targets for the length of judicial proceedings? |
| Psychological pressure and stress | I don’t have enough time to study the cases properly  
I am always well prepared for hearings  
I often feel forced to excessively simplify the reasoning of my written judgments, in order to dispose enough proceedings  
There are significant consequences for judges if the targets for the length of proceedings are not met  
I think that the focus on efficiency is compromising the quality of my decisions  
The standardisation of procedures affects my autonomy  
The control /pressure by the court/section president affects my independence in a negative way  
Judges’ failure to meet targets has significant consequences for the court  
Judges’ failure to meet targets has significant consequences for the judge  
Performance targets are too ambitious  
Section/court presidents check that all judges meet performance targets  
My colleagues monitor that everyone fulfils the performance targets  
It is difficult to meet the targets for the timeframes of judicial proceedings set by the court  
The court/section president takes action if the length of proceeding is about to exceed the target  
There are significant consequences for judges if the targets for the length of proceedings are not met  
There are standard instructions/practices/rules to schedule hearings  
I think there’s too much emphasis/pressure on efficiency  
On a scale from 1 to 5 (where 5 is the maximum level), how do you rate the sustainability of your workload - considering an average of the last two years  
On a scale from 1 to 5 (where 5 is the maximum level), how do you rate your work-related stress?  
To what extent do you think the following elements put pressure on you, leading to work-related stress?  
- Pressure from the president (section or court) on the organisation of my work  
- Pressure from colleagues  
- Pressure from the public opinion  
- Pressure from the Council for the Judiciary  
- Pressure from the Ministry  
- Pressure from parties and/or lawyers  
- Lack of adequate working space  
- Lack of adequate ICT tools (hardware, software)  
- Lack of adequate administrative support (clerks, judicial assistants etc)  
- Difficulty of work-life balance |
| Stress reducers | I am autonomous in deciding which cases to deal with first  
|                | I am autonomous in scheduling my hearings  
|                | I am autonomous in deciding how to organise my work  
|                | Judge’s opinion is taken into account in setting the performance targets  
|                | Dialogue between president and judge is the main way to address individual performance issues  
|                | I give the needed attention to each case, even if this may increase the time of disposition  
|                | I give the needed attention to each case, even if this may decrease the court’s efficiency  
|                | On a scale from 1 to 5 (where 5 is the maximum level), how do you rate your autonomy in organising your work?  
|                | Which factors help you to relieve your work-related stress?  
|                | - support from administrative staff  
|                | - support from colleagues  
|                | - support from the president  
|                | - Strong motivation  
|                | - Perception of my role as a judge in the community  
|                | - Autonomy / freedom in managing my time and priorities  
|                | - Level of salary  |
Appendix 2. Results of ordinal regression analysis

| Ordinal regression with work-related stress as dependent variable: parameter estimates |
|-----------------------------------------------|-------|-----------|-----------|
|                                              | Estimate | Std. Error | Sig.       | Lower Bound | Upper Bound |
| Intercept stress: very low                   | -4.791   | 0.388      | 0.000      | -5.551      | -4.032      |
| Intercept stress: low                        | -2.869   | 0.301      | 0.000      | -3.480      | -2.279      |
| Intercept stress: high                       | -0.690   | 0.269      | 0.010      | -1.218      | -0.162      |
| Intercept stress: very high                  | 1.917    | 0.287      | 0.000      | 1.354       | 2.479       |
| Gender: male                                 | -0.626   | 0.150      | 0.000      | -0.919      | -0.333      |
| Hours worked < 35                            | -1.033   | 0.330      | 0.002      | -1.680      | -0.387      |
| Hours worked 35 - 40                         | -0.326   | 0.214      | 0.127      | -0.745      | 0.093       |
| Hours worked 45 - 50                         | 0.491    | 0.193      | 0.011      | 0.113       | 0.869       |
| Hours worked > 50                            | 1.552    | 0.236      | 0.000      | 1.090       | 2.013       |
| Country Finland                              | -0.707   | 0.251      | 0.005      | -1.199      | -0.215      |
| Country Netherlands                          | -0.361   | 0.184      | 0.049      | -0.722      | -0.001      |
| Not enough time: strongly disagree           | -1.220   | 0.410      | 0.003      | -2.023      | -0.417      |
| Not enough time: disagree                    | -0.376   | 0.199      | 0.059      | -0.768      | 0.014       |
| Not enough time: agree                       | 0.657    | 0.206      | 0.001      | 0.253       | 1.060       |
| Not enough time: strongly agree              | 2.118    | 0.342      | 0.000      | 1.448       | 2.788       |
| Sign. consequences failure: strongly disagree| -0.084   | 0.275      | 0.758      | -0.623      | 0.454       |
| Sign. consequences failure: disagree         | -0.022   | 0.176      | 0.901      | -0.368      | 0.324       |
| Sign. consequences failure: agree            | 0.152    | 0.216      | 0.482      | -0.272      | 0.576       |
| Sign. consequences failure: strongly agree   | 1.031    | 0.394      | 0.009      | 0.258       | 1.803       |
| Pressure from parties/lawyers: very low      | -1.065   | 0.279      | 0.000      | -1.612      | -0.518      |
| Pressure from parties/lawyers: low           | -0.411   | 0.165      | 0.013      | -0.735      | -0.088      |
| Pressure from parties/lawyers: very high     | -0.152   | 0.261      | 0.560      | -0.663      | 0.359       |
| Pressure from lack administrative staff: very low| -0.213   | 0.273      | 0.434      | -0.748      | 0.321       |
| Pressure from lack administrative staff: low | -0.284   | 0.183      | 0.150      | -0.623      | 0.096       |
| Pressure from lack administrative staff: very high| 0.706    | 0.196      | 0.000      | 0.325       | 1.088       |