Tools, methods, and some caveats to analyze the Brazilian Judiciary Performance data

Costa Louro, Alamir; Zanquetto Filho, Hélio; Romão dos Santos, Washington; Moll Brandão, Marcelo
Tools, methods, and some caveats to analyze the Brazilian Judiciary Performance data
Administração Pública e Gestão Social, vol. 13, núm. 1, 2021
Universidade Federal de Viçosa, Brasil
Disponible en: http://www.redalyc.org/articulo.oa?id=351564966008

Tools, methods, and some caveats to analyze the Brazilian Judiciary Performance data

Ferramentas, métodos e algumas ressalvas para analisar os dados de desempenho do Judiciário brasileiro

Herramientas, métodos y algunas advertencias para analizar los datos de desempeño del Poder Judicial brasileño.

Alamir Costa Louro
Universidade Federal do Espírito Santo, Brasil
alamirlouro@gmail.com

Hélio Zanquetto Filho
Universidade Federal do Espírito Santo, Brasil
zanquetto@gmail.com

Washington Romão dos Santos
Universidade Federal do Espírito Santo, Brasil
washington.santos@ufes.br

Marcelo Moll Brandão
Universidade Federal do Espírito Santo, Brasil
mollmkt@gmail.com

Resumo:

Objetivo da pesquisa: O artigo discute os antecedentes do desempenho do Judiciário no Brasil, baseando-se no modelo de Dimitrova-Grajzl, Grajzl, Sustersic e Zajc (2012) sobre os tribunais eslovenos.

Enquadramento teórico: O conteúdo substantivo do presente modelo, no estudo, considerou discussões de trabalho basilar anterior e sua operacionalização está anexada aos indicadores do Conselho Nacional de Justiça (CNJ) de 2016, por meio do relatório "Justiça em Números" do Conselho Nacional de Justiça.

Resultados: Os resultados quantitativos mostram os perfis da força de trabalho em detalhes, a relação linear e o efeito moderador da carga de trabalho como não significativo, enquanto seu efeito quadrático enriquece a discussão sobre a produtividade.

Originalidade: Os resultados e o caminho metodológico e as advertências sobre os dados são oportunidades para trabalhos futuros para explicar as ambiguidades quando o desempenho é comparado usando Dimitrova-Grajzl e modelos de trabalho atuais.

Contribuições teóricas e práticas: Dados do CNJ precisam ser utilizados com parcimônia e diversas técnicas aqui apresentadas ajudam nesse sentido.

Palavras-chave: Desempenho Judiciário, OLS, PLS-SEM, IPMA, Não linear.

Abstract:

Research Objective: This article discusses the judiciary's performance antecedents in Brazil. It based its model on the work of Dimitrova-Grajzl, Grajzl, Sustersic and Zajc (2012) about the Slovenian courts.

Theoretical framework: The substantive content of the present model received simplified discussions because it is based on previous basilar work and its operationalization is attached to CNJ, the Brazilian National Council of Justice 2016 indicators, through the "Justice in Numbers" report by the referred council.

Design: The constructs performance, human capital, expenditures, investment in new technologies, and caseload are measured using OLS, PLS-SEM, IPMA, and quadratic regression.

Findings: The quantitative results show workforce profiles in details, the linear relationship, and the moderating effect of workload as not significant, while its quadratic effect enriches the discussion about productivity.

Originality: the results and the methodological path and data caveats are opportunities for future works to explain the ambiguities when performance is compared using Dimitrova-Grajzl's and present work models.

Theoretical and practical contributions: CNJ data must be used sparingly and several techniques presented here help in this regard.

Keywords: Judiciary Performance, OLS, PLS-SEM, IPMA, Non-Linear.

Resumen:

Objetivo de la investigatión: Analizar los antecedentes de desempeño del poder judicial en Brasil. Basó su modelo en el trabajo de Dimitrova-Grajzl, Grajzl, Sustersic y Zajc (2012) sobre los tribunales eslovenos.
Marco Teórico: El contenido sustantivo del presente modelo recibió discusiones simplificadas porque se basa en un trabajo básico previo y su operacionalización se adjunta a los indicadores del CNJ, el Consejo Nacional de Justicia de Brasil (2016), a través de su informe "Justicia en números".
Metodología: Los constructos el rendimiento, el capital humano, los gastos, la inversión en nuevas tecnologías y la carga de casos se miden utilizando OLS, PLS-SEM, IPMA y regresión cuadrática.
Hallazgos: Los resultados cuantitativos muestran los perfiles de la fuerza laboral en detalle, la relación lineal y el efecto moderador de la carga de trabajo como no significativos, mientras que su efecto cuadrático enriquece la discusión sobre productividad.
Originalidad: Los resultados y la trayectoria metodológica y las advertencias de los datos son oportunidades para que los trabajos futuros expliquen las ambigüedades cuando se compara el desempeño utilizando los modelos de trabajo de Dimitrova-Grajzl y el presente.
Contribuciones teóricas y prácticas: Los datos del CNJ deben usarse con moderación y varias técnicas presentadas aquí ayudan en este sentido.

Palabras clave: Desempeño del Poder Judicial, OLS, PLS-SEM, IPMA, No lineal.

1 Introduction

The purpose of the present paper in testing the relationships between workforce and workload with Judicial Performance is in agreement with the goal of Dimitrova-Grajzl *et al.* (2012) work, which is considered a cornerstone for the definition of the constructs used. The epistemological positioning of the present authors is to recognize the existence of universal theories, but the work pays attention to the different contexts that Brazil and Slovenia have, not generalizing its results but raising analysis and caveats, to data and method, which can be seen as opportunities for future studies. The field that studies Judicial Performance is assumed as initial (Guimaraes, Gomes, & Garrido, 2018) and exploratory. Thus a data-driven emergent theory (Jaccard & Jacoby, 2010) is pointed out as a better approach for understanding the phenomenon.

In the international literature review conducted by Dimitrova-Grajzl *et al.* (2012), they pointed out the ambiguity in results about the court output, Judicial Performance, and its antecedents, caseload, and workforce. It is important to emphasize that all constructs were measured differently in present work. Notwithstanding, the issue of the endogeneity of these constructs, essential to the cornerstone work, is not exhausted discussed in our study, assumed as exploratory. Additional meanings are possible about this constructs, as risks (Capaldo, Costantino, Pellegrino, & Rippa, 2018), for example. However, we assume a different measure but the same meaning of the cornerstone work.

We consider that any other analysis of empirical tests without contextual assumptions can lead to a wrong application of conceptual validity of the constructs previously defined by other authors, taking away the original explanatory power of the previous work, but aside from all this research delimitation and the apparent limitations, we assume that this doesn't change the importance of empirical verification/validation of another model, as we did in the first step of present paper.

The national literature brings out an inefficiency image of the Brazilian Judiciary (Sadek, 2004; Louro, Santos, & Zanquetto, 2017; Guimaraes, Gomes, & Guarido, 2018). However, the international studies of efficiency are more often than inefficiency studies (Mitsopoulos & Pelagidis, 2007; Falavigna, Ippoliti, Manello, & Ramello, 2015; Teixeira, Bigotte, Repolho, & Antunes, 2018). A possible explanation of this bad image may come from the national judicial system, as a codified law, that may not include the constant social changes. However, it brings it out in more critical thought. This such view is rejected as it seems to be

ethnocentric rationalism established by modern thinking and "accepted" in developing countries with "less juridical tradition."

The pessimist view makes "unavoidable" the inefficiency of the Brazilian legal system, but getting out of the ideographs the present study is based only on nomothetic, and the positivist science approaches, exactly to contradict some take for granted arguments. Another view of Judicial Performance is a balance between efficiency and quality of these courts, functioning as one characteristic which necessarily excludes the other. Again the present work understands that this is a rhetorical proposition, and is created exactly by the "local judicial culture" about expectations and subjective perceptions, what is different from one environment to other (Ng, 2011).

Some literature about Judiciary in Brazil uses the "Justice in Numbers" report of National Justice Council (Conselho Nacional de Justiça [CNJ])(Nogueira, Oliveira, Vasconcelos, & Oliveira, 2012; Gomes, Guimaraes, & Akutsu, 2017; Louro, Santos, & Zanquetto, 2017), but few take additional methodological steps for robust results or even use weaker methods. We point out some caveats to this dataset, and we raise some methodological issues to improve future studies in Brazil.

The prominent contribution of the present paper is our theoretical model tests using workload and workforce as antecedents of Judicial Performance. Additionally, we give steps to use the CNJ data for providing an excellent view of the possibilities for future works. Methodologically, we used structured Equation Modeling (SEM) with Partial Least Square (PLS) algorithm, Ordinary Least Square (OLS) in IBM SPSS and in R (R Core Team, 2018). Synthetically, the paper showed that in Brazil the role of workload needs profound understanding. Additionally, it showed that workforce construct is a robust predictor and its profiles have different effects on Performance, especially when it is analyzed considering money spent.

2 THEORETICAL REFERENCE AND DEVELOPMENT OF THE MODEL

The substantive content of the present model received simplified discussions because it is based on previous basilar work and its operationalization is attached to CNJ (2016) indicators, through the "Justice in Numbers" report of National Justice Council of Brazil.

2.1 Workload

According to Gomes, Guimaraes, and Akutsu (2017), Brazilian courts have faced an imbalance between the demand versus capacity after National last Constitution formalization. This imbalance increased congestion rate, an index defined by the National Council of Justice as a relation between new cases and the number of processes filed (CNJ, 2016), and directly related to workload, that is a backlog or the result of congestion.

According to Dimitrova-Grajzl *et al.* (2012), the pressure caused by the increase of judge's workload has an impact on performance, and for these authors, the workload is the number of processes not filed from the previous years and the new processes of the current year. This is the same definition from National Council of Justice direct indicator (CNJ, 2016).

2.2 Workforce

The human resources that assist the judges can be classified into servers, Commissioners, outsourced and trainees, it is assumed that the work of all they can directly affect the performance of the Courts. At this point, the research differs from the approach of Dimitrova-Grajzl *et al.* (2012) which only considered the number of judges. Louro, Santos, & Zanquetto, 2017, Gomes, Alves, and Silva (2018), already discussed the possibility of using the different profiles as a unique Human Capital or workforce construct, and we advance

its analysis with additional tests to show the different importance and performance of these profiles. Then, the workforce was operationalized by the number of judges plus officials that help the judicial work done by the judges: commissioned employees, interns, servants, and outsourced. These officials assist judges in many judicial activities, for example, preparation of legal opinions, holding of hearings, mediation between litigants, reports, among others (Gomes *et al.*, 2018).

2.3 Productivity

The CNJ report specified its productivity measurement, or Judicial Performance, as the "Number of Processes filed" indicator (CNJ, 2016). Productivity is interchangeable with Judicial Performance as an assumption here, of course, a simplification because Courts goals can be diverse (Sadek, 2004). Then, just like Gomes *et al.* (2018), the number of the file processed was used to define the measured productivity construct. An important caveat here is that Productivity should be a ratio related to number of professionals, but process filed are worked by different profiles then we adopt the number, not the rate.

2.4 The relationship between Workforce and Productivity

Nogueira, Oliveira, Vasconcelos and Oliveira (2012) affirm that there has been an increase in productivity of the Brazilian Courts of Justice, but they also say that there is no consensus on what are their most important antecedents. Researchers of judiciary productivity has examined the role of judges and possible reflections on productivity; on the other hand, as emphasized by Ng (2011) inevitably changes in productivity imply bringing judges to a new management field, becoming productivity and backlog responsible, so it's necessary to evaluate issues about accountability shared with courts or even with National Council of Justice.

However, the accountability issue embraces a subjectivity that is out of the scope in this paper, on the other hand, the present work focus on the backlog, the workload related to old cases, the backlog management is understood as workforce time spent not finishing processes then do not count to productivity.

Results of studies (Gomes *et al.*, 2018; Louro, Santos, & Zanquetto, 2017) indicate that there is a positive relationship between the workforce size and courts productivity, so the courts with the largest number of assistants, servants, judges and outsourced, are the ones that present the best results in terms of productivity.

Although it is simple to think that with more people we can have more processes filed, it is not obvious, because other variables can influence the Performance explanation. In this study we consider that the relationship between workforce and productivity is positive, that is, the larger the workforce available, the greater the number of processes filed. Therefore, the present work begins with the primary conceptual hypothesis:

Hypothesis 1. (H.) Workforce (Judges, Servers, Commissioners, Outsourcers, and Interns) has a direct positive effect on Court Productivity.

2.5 The relationship between Judicial Workload and Productivity

For Dimitrova-Grajzl *et al.* (2012) there is an ambiguous relationship between workload and Performance. Primarily thinking, judges should avoid backlogs but when already there is a considerable backlog, its management looks like to take most of the staff time. Workload behavior in the words of Dimitrova-Grajzl *et al.* (2012):

> "The magnitude of the impact of caseload on the number of resolved cases, therefore, depends on the interplay of these two effects: when the incentive effect is weakened by the congestion effect, we would expect that the resolved cases will

be inelastic with respect to caseload or, hypothetically, even negatively related to caseload. In contrast, when the incentive effect strongly dominates the congestion effect, we would expect the number of resolved cases to be elastic with respect to increases in caseload."(p. 22)

In Brazil, the incentive to finish the backlog, the Workload, the result of congestion, is given by CNJ policies. Back and agreeing with Ng (2011) critics, the "due process" and "access to justice" by a reasonable part of the population seems to be in opposite position in the balance of the CNJ, the institution that demands for court productivity but doesn't show the same attention to Courts Workload, the focus of the present work.

The judicial congestion is essential in the context of this study because it sets the increasing backlog that must be executed by the courts (Procopiuck, 2018), which directly involves the judicial staff and capacity of courts to process and manage the cases. In the next hypothesis, we anticipate that the workload acts linearly to increase the performance, and it also has a quadratic effect that helps us to understand better the workload construct.

A larger support team creates a better work environment for the judges, which tends to increase the speed and productivity of the courts (Gomes *et al.* 2018). However, the present work understands that the linear relationship is not clear to Dimitrova-Grajzl *et al.* (2012) and it is necessary to confirm this hypothesis:

Hypothesis 2. (H_{2a}) Workload has a direct and negative effect on Productivity.

Differently, from the apparent previous relationships based on Dimitrova-Grajzl *et al.* (2012) work, the present work needs to test other possibilities of connections between the workload and Productivity. In the beginning, the workload and the incentives of CNJ can improve the productivity, but the greater the workload in the judiciary, without the workforce in the same proportion it can reduce the productivity, i.e., a U-inverted shape is thought to represent the quadratic relationship. Therefore, we tested the following hypothesis that Dimitrova-Grajzl *et al.* (2012) did not test:

Hypothesis 2. (H_{2b}) Workload has a quadratic effect on Productivity.

Differently, from the cornerstone work, we already tested the moderation of Workload, to understand if it can be a contextual effect in the relationship between workload and productivity. The backlog management is essential in the context of this study because it sets the volume of work that must be performed in courts, which directly involves the judicial staff. The workload acts as a moderating variable, again with a negative effect. In other words, it is expected that the relationship between workforce and productivity is mitigated by the number of processes not filed in courts (workload).

Hypothesis 3. (H.) Workload has a moderating effect decreasing the causal relationship between Workforce and Productivity.

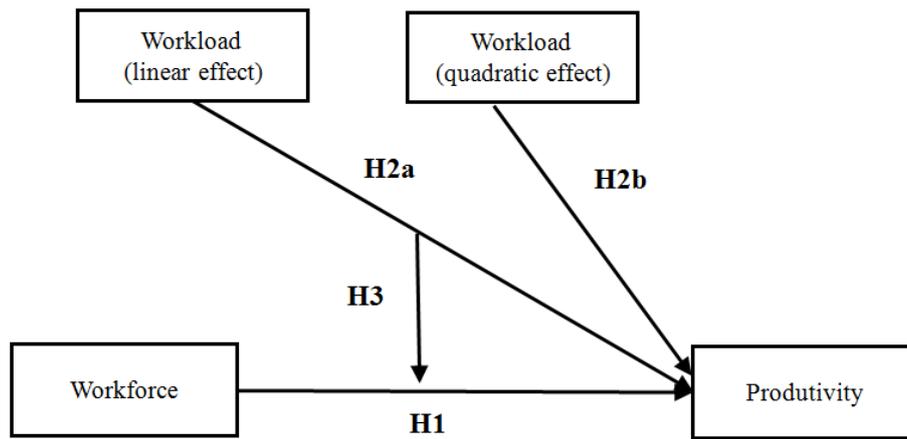The hypotheses result in the theoretical model presented in Fig. 1.

Fig. 1 - Theoretical Model
Source: Prepared by the authors (2020).

An additional analysis was executed to boost the understanding of the Quadratic Effect of H2b, and all caveats necessaries were given. Additionally, an importance-Performance Map Analysis (IPMA) was performed to improve the knowledge about the role of workforce profiles in the sample of Brazil.

3 Methodology

As can be seen in the model development, the constructs are taken for granted because it is a known secondary data that fits well the discussion with Dimitrova-Grajzl et al. (2012), we assume this as a limitation, and we raise some methodological caveats about the dataset and the constructs operationalization.

3.1 Data and measurements

The present work has a quantitative nature and is characterized as exploratory because we understand all the topics as early stage theoretical constructs in an emerging research field (Guimaraes, Gomes, & Guarido, 2018). In this case, it is widespread for the researcher to face with unobserved heterogeneity, omitted variable, omitted selection, simultaneity, measurement errors to cite some caveats that we will expose.

It was used the secondary data from the "Justice in Numbers" report, totaling (n = 527) records to test the hypotheses, quantitative data from State Courts of Justice, and Federal Courts (Federal Regional Courts, Regional Electoral Courts, and Regional Labor Courts) that provided information for seven years, 2009 until 2015. The years of 2009 and 2010 do not have data for Regional Electoral Courts. From this amount were found eleven records of the "outsourced" variable as missing data, representing less than 5% of the population, which according to Hair, Black, Babin, Anderson and Tatham (2009) can be classified as lost and non-ignorable data, as randomized data was chosen as the substitution by the respective mean.

3.2 Data and methods Caveats

Table 1 shows the descriptive statistics that indicate data non-normality, showing the referenced codes by the "Justice in Numbers" report and the variables that appear in the measurement model shown in Fig.2, this care is vital for the reproducibility of the work.

Table 1- Non-Normality Indication

Descriptive Statistics

	Mean	Std. Deviation	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
Intern (tfauxe)	661.324	1266.392	4.337	0.106	23.847	0.212
Outsourced (tfauxt)	745.992	1243.307	4.167	0.106	22.548	0.212
Servants (serv)	3391.879	5730.757	5.402	0.106	35.412	0.212
Commitioners (cc)	559.875	1366.908	6.945	0.106	54.404	0.212
Judges (mag)	246.110	328.708	4.549	0.106	26.970	0.212
Workload	3036.994	2776.186	1.549	0.106	3.819	0.212
Processes Filed(tbaix)	351235.738	709470.424	4.847	0.106	28.950	0.212

Source: secondary CNJ data (using IBM SPSS).

Table 1 indicates non-normality because of kurtosis and skewness are not inside the interval from -1 and 1 (Hair et al., 2009). To confirm the univariate result, we executed Shapiro-Wilks and Kolmogorov-Smirnov tests failing to reject the hypothesis of normality for all variables (Hair et al., 2009). The normality condition is one of the requisites for Ordinary Least Square (OLS) regressions and covariance-based approaches (CB-SEM) (Hair, Hult, Ringle, & Sarstedt, 2017). Additionally, we understand the work as exploratory by start, that is why we used PLS-SEM. Another argument is that SmartPLS (Ringle, Wende, & Becker, 2014) software have different tools that gave us valuable information like the Importance-Performance Map Analysis (IPMA), offering information on the relative importance of constructs in explaining other constructs in the structural model (Hair et al.,2017).

The first data's caveat is the quality of the dataset about measurement error. Certainly, a report that receives data from all Courts of Brazil (CNJ, 2016), that have different Information Systems have fragilities of measurement its indexes. Traditional multivariate procedures are incapable of assessing measurement error, and SEM provides explicit estimates of these error variance parameters (Hair et al.,2017).

A common mistake, in any non-experimental work, is the sample selection bias, this is due to a nonrandom sample (Bettis, Gambardella, Helfat, & Mitchell, 2014). The present dataset does not permit experimental methods because the CNJ (2016) report shows data from all Courts in Brazil, but we tested differences between groups to improve it methodologically and check the quality of the dataset.

In secondary data a mistake is about data heterogeneity, that means the existence of different behaviors between groups. A multi-group analysis was executed in SmartPLS software, based on the nonparametric equivalence analysis technique called PLS-MGA (Partial Least Square - Multi-Group Analysis), considered an original extension of Henseler et al. (2009) MGA method. It was defined as groups for the test the existence of States Courts and Federal Courts.

The PLS-MGA test revealed that all p-values of the difference between the groups (i.e., group 1 States - group 2 Federal) are not significant, indicating, therefore, the possibility of pooling the data (Hair et al., 2017). Additionally, to test measurement invariance, we used the MICOM process (Henseler, Ringle, & Sarstedt, 2016) executing the permutation algorithm with 5000 permutations confirming no significance and then measure invariance. A caveat is necessary because not all groups possible were thought. For future

studies, it is important to say that the not found heterogeneity is specific to our sample choice of variables and we affirm that with other variables it is not true. Thus it needs more tests.

3.3 Step 1- Structural Equation Modeling Analysis

The data were investigated using the structural equation modeling (SEM) that was used to test the hypotheses of the model. According to Hair et al. (2009), a sample without good normal distribution data and the fact that the model has five latent variables, and therefore, several interrelated dependence relationships. This led to the use of SEM, thus SmartPLS software in its version 3.2.4 was chosen, which brings the PLS statistical method, Partial least squares, which requires relatively few registers/cases, the parameter pointed out by Hair et al. (2017), in the worst case, as 10 times the highest number of arrows pointed at any latent variable. Then in the present study would be 50 records, then as the sample consisted of 527 records, it was possible to satisfy this condition.

The PLS algorithm was executed with the default values following Hair et al. (2017), the Quadratic hypothesis was executed in R, results are shown in Fig.2.
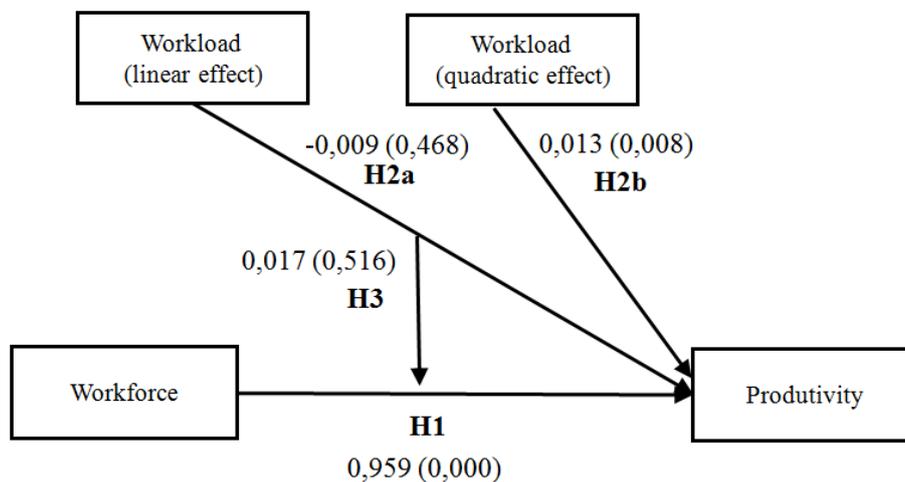

Fig. 2 - Results for direct and moderating hypotheses (SmartPLS), Quadratic (R)
Source: Prepared by the authors (2020).

The unique construct with more than one variable, or National Council of Justice report indicator, was considered reflexive by its definition. Validity and reliability results are shown in Table 2, which presents Cronbach's Alpha and composite reliability, greater than 0.7 and the Average Variance Extracted (AVE), greater than 0.5.

Still referring to convergent validity, it was found that the external loads, not shown here, are all higher than 0.924, thereby, according to what Hair et al. (2017) affirm that they should be eliminated when it's less than 0.4 and evaluated if less than 0.7.

Table 2 - Reflective Measurement Model

Construct	Cronbach's Alpha	rho_A	Composite Reliability	AVE
Workforce	0.9761	0.978	0.982	0.915

Source: secondary CNJ data.

Still talking about the measurement model, analyzing the discriminant validity, according to the Fornell-Larcker criterion, the square root of the AVE of the reflexive constructs must be greater than the loads of the other constructs. It was not found a problem between Workforce (0.956) and Productivity (0.968). However, we understand that the content validity of the Workforce and Productivity constructs do not generate theoretical doubts. The same behavior is repeated in the cross-load test, but we hold to the definitions of both constructs that are based on the National Council of Justice's definitions. According to Hair et al. (2017), the first step for the evaluation of the structural model is to evaluate collinearity using the VIF indicator, which parameter must be less than 5, the highest result found was 1.173. The second step is to evaluate the path coefficients, and for this the Bootstrapping procedure was executed with 5000 subsamples with the option "no sign changes", resulting in Table 3 that shows no significance for the path between "Workload" and "Productivity", neither for the path between the mediator "Workforce X Workload" and "Productivity", the other path followed the parameters by Hair et al. (2017), t statistic > 1,96 or p-value <0.05.

Table 3 - Bootstrapping

Path coefficients	Original Sample (O)	Sample Mean (M)	Standard Deviation (STDEV)	T Statistics (|O/STDEV|)	P-Values
Workforce ->Productivity	0.959	0.963	0.022	44.590	0.000***
Workforce X Workload ->Productivity	0.017	0.015	0.026	0.650	0.516
Workload ->Productivity	-0.009	-0.010	0.012	0.726	0.468

Source: secondary CNJ data (using Smart PLS).
* p < .05, ** p < .01, *** p < .001.

The third step is to evaluate the coefficient of determination that measures the predictive power of the model, the results show substantial value (R Square and R Square Adjusted of .938), according to Hair et al. (2017), for Productivity construct. Step four attempts to measure the size of the f^2 effect that evaluates if any omitted constructs generate a substantive impact on endogenous constructs. In the case, just the Workforce (17.271) construct is relevant to Productivity.

Table 4 - Blindfolding - Cross-validated Redundancy

	SSO	SSE	Q² (=1−SE/SSO)
Productivity	527.000	44.405	0.916
Workforce	2.635.000	2.635.000	
Workforce X Workload	527.000	527.000	
Workload	527.000	527.000	

Source: secondary CNJ data (using Smart PLS).

In the fifth step, as shown in Table 4, the predictive relevance is evaluated using the Blindfolding algorithm using the default configuration, e.g., omission distance equal to seven. This resulted in a Q^2 that represents great relevance for the Productivity construct (0.916), using the parameter higher than 0.35 according to Hair et al. (2017).

Another test performed was to analyze whether Workload has a moderator effect, however, after another bootstrapping using the two-stage method, including an interaction term, the p-value of the path coefficient

between the moderator effect Workload in the causal relationship between the Workforce exogenous construct and the Productivity endogenous construct resulted in non-significance 0.516, rejecting H3.

A parallel discussion about Dimitrova-Grajzl et al. (2012) work is that they pointed to a possibility of reverse causality, which was tested in our sample, and confirmed statistically. However, we affirm that this conjecture does not make sense because of the causal claim temporality condition, i.e., to have a causal relationship, the independent variable should exist before the dependent variable, and the workforce exists to fill processes, not the opposite. Additional, the non-significance of Workload as a direct measure or even as a moderator in the relationship of Workforce with Productivity means, in our sample, the statistical inviability of simultaneity, but gives an opportunity for future works.

We raise an essential caveat here because in an emerging field indeed omitted variables and simultaneity problems are hard to find, and it's an essential matter for confirmatory research, and as we said some papers use the wrong method and few limitations (Antonakis, Bendahan, Jacquart, & Lalive, 2010) to claim causality.

After this conclusion, it is possible to exclude the Workload as direct antecedent of Performance. The result that makes viable a multiple regression, OLS, where the dependent variable is Processes Filed and the independent variables are the various labor profiles. Of course, this is possible only ignoring the caveats that we presented, but for exploratory research in emergent theories, the pros are greater than cons.

3.4 Step 2- Importance-Performance Matrix Analysis (IPMA)

The last test performed in SmartPLS is the Importance-Performance Matrix Analysis (IPMA) to verify which exogenous variable are most relevant to the relationship between performance and importance (Fig. 3). This analysis is a contribution to managers and legislators. Indeed it has some limitations about the profile information but gives insightful perceptions about how to invest in Human Capital.

IPMA shows the Importance (X-axis) as the structural model total effects in shaping the target (Productivity) construct, versus the Performance (Y-axis) as the average latent variables (profiles) scores. In the present case, we balanced the Importance using different weights by profile. We did two different analyses, considering all profile equally, and considering the profiles versus how much the government spends with it (see Fig. 3).
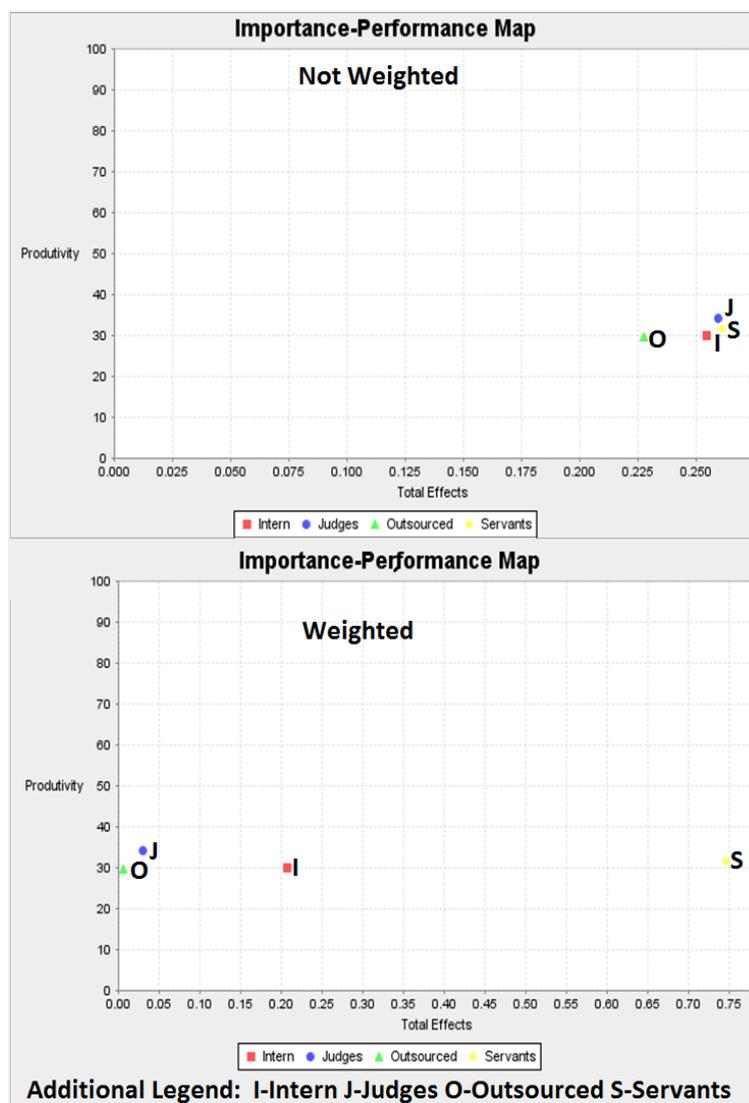
Fig. 3 - Importance-Performance Matrix Analysis (IPMA)
Source: Prepared by the authors (2020).

For the second analysis (bottom side), we assumed different weights between the profiles according to the information of average expenditure. We used the simple mean of CNJ (2016) indicators G10a (Judge), G10b (Servant), G10c (Outsourced) and G10d (Intern). Unfortunately, there is no indicator to the Commissioners average expenditure, that is why it is out of both analyses (see Table 5 for profile mean information).

Table 5 - Weight Rule

Profile	Average Expenditure	Intern Baseline	Weight (1/Baseline)
G10a (Judge)	R$430352.00	R$51.32	0,01949
G10b (Servant)	R$157003.00	R$18.72	0,05341
G10c (Outsourced)	R$33607.90	R$4.01	0,24952
G10d (Intern)	R$8385.86	R$1.00	1,00000

Source: secondary CNJ data.

The upper side of Fig. 3 (not weighted) result shows that all the profiles have nearly similar importance(effect) and similar performance(mean). On the other hand, the bottom side of Fig. 3 (weighted) shows that investments on Servants and Interns are the most important considering the money spent for total effect on Productivity.

Some considerations are needed here, the Intern profile have relatively low cost, but of course, this analysis does not show any aspect about Quality, it just analyses the number of processes filed. Another important consideration is that despite the low total effect of Judges, they are required by law as a critical path in the process.

In Brazil, judges' assistants, the other profiles, play an important role in judicial production, as they often develop specific legal activities for judges, such as the preparation of judgments and decisions, agreements between the parties and the organization of cases in the judicial unit (Gomes et al., 2018). In both analyses, the outsourced profile has the least importance to Judiciary Performance, and maybe it can be explained because this profile has more administrative work, that is, activities that would have no direct relation to the judicial work done by the judges.

3.5. Step 3- Ordinary Least Squares (OLS) analysis

The IPMA analysis did not count with commissioners, but the OLS can analyze all profiles. But before the regression execution, we present underused essential information in exploratory studies, the matrix of correlation (Bettis et al., 2014), Table 6 shows the existence of some relevant correlations that encourages the study continuation.

Table 7 - Simple Regression R^2

Model 1 Summary – With Influencers

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.977a	.955	.955	153731.039440

Model 2 Summary – Without Cook's D Influencers

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
2	.976a	.953	.953	142698.922693

a. Predictors: (Constant), Zscore(Servants), Zscore(Outsourced), Zscore(Commissioners), Zscore(Intern), Zscore(Judges)

Source: secondary CNJ data (using IBM SPSS).

As said earlier, there is not only the non-normality problem in CNJ dataset, another caveat in some OLS studies using the same or similar datasets is related to the presence of influencers, and this process looks like to be ignored or even used as p-value hacking practices (Head, Holman, Lanfear, Kahn, & Jennions, 2015).

Cook's D and leverage(hat-value) influencers can be seen in Fig. 4, it was used to separate the datasets of Model 1 and 2. The cutoff parameter used to separate the datasets was 1 (Cohen, Cohen, West, & Aiken, 2003).

It is not robust to use the CNJ (2016) dataset without the evaluation of influencers and outliers. Using "Car" package (Fox & Weisberg, 2011) in R (R Core Team, 2018) was found three possible outliers (registers 26,82,490 with p-value and Bonferroni adjusted p-value <.0001) but just two influencers (registers 26,82).
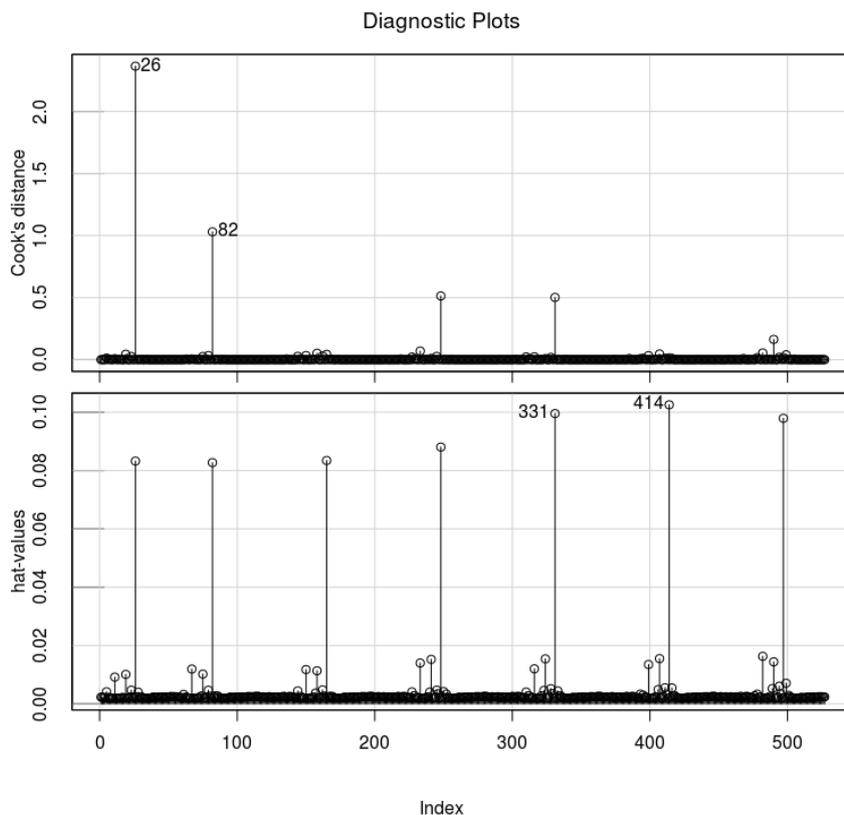
Fig 4 - Cook's D and hat-values (using R and "car"package)
Source: Prepared by the authors (2020).

Table 8 shows the differences of significance and coefficients of the slopes when influencers are manipulated; the regressors were standardized to z-scores according to Cohen et al. (2003).

Aside from the outliers and influencers presence caveat in the dataset of CNJ (2016), some new discussions arise from the last regression result (without influencers in bottom sided of Table 8, i.e., the best dataset for regression).

Table 8 - Simple Regression Intercept and Coefficients

Model 1 - Coefficients^a With Influencers

		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
Model		B	Std. Error	Beta		
1	(Constant)	354965.202	6849.557		51.823	.000***
	Zscore (Commissioners)	114093.750	18916.560	-.161	-6.031	.000***
	Zscore (Intern)	184810.685	20874.680	.258	8.853	.000***
	Zscore (Judges)	88072.067	28919.636	.124	3.045	.002**
	Zscore (Outsourced)	42387.970	16830.764	.059	2.518	.012*
	Zscore (Servants)	500638.568	35065.994	.706	14.277	.000***

Model 2 - Coefficients^a- Without Cook's D Influencers

		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
Model		B	Std. Error	Beta		
2	(Constant)	353938.870	6374.412		55.525	.000***
	Zscore (Commissioners)	107195.300	17722.867	**-.151**	-6.048	.000***
	Zscore (Intern)	181322.553	19403.751	**.263**	9.345	.000***
	Zscore (Judges)	107329.044	26929.147	.151	3.986	.000***
	Zscore (Outsourced)	66761.671	17985.873	.101	3.712	.000***
	Zscore (Servants)	448589.798	35299.713	**.623**	12.708	.000***

a. Dependent Variable: Processes Filed

Source: secondary CNJ data (using IBM SPSS).
$^* p < .05, ^{**} p < .01, ^{***} p < .001.$

The impact level of each profile, highlighting the coefficients of Servants and Interns, and the specific negative result of the commissioners coefficient, raise other critical discussions to be clarified in future works.

3.6 Step 4 - Quadratic analysis

Using "ggplot" package in R, we executed the scatterplot of Performance versus Workload (see Fig. 5). It demonstrate the data behavior, level of dispersion and effect between the two variables of quantitative nature, Workload and Productivity. This Fig. 5 shows us clues, U-inverted shaped of the bootstrap envelop, shadow, revealing a not linear behavior, that can mean that the workload increase the productivity in the beginning and after a point, the productivity remains inelastic as Dimitrova-Grajzl et al. (2012) predicted but not tested, or as can be seen, it decreases productivity.
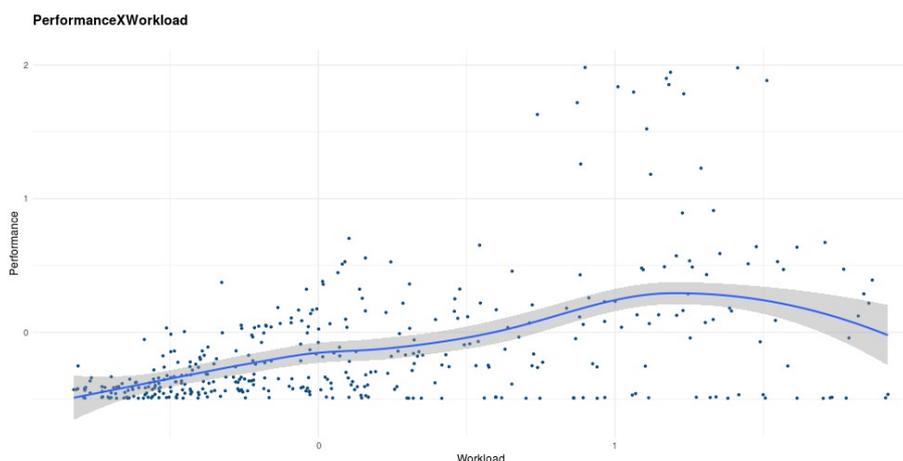
Fig. 5 - Scatterplot (using R and "ggplot"package).
Source: Prepared by the authors (2020).

The analysis of this behavior will not be exhausted by the present work, but we give a step ahead. The regressions for Quadratic analysis (see Table 9 and 10) were executed in R to take advantage of the simple model comparison using F statistics of the package "stats", a core part of R (R Core Team, 2018). The results of analyzes show that the model 2 (quadratic analysis) presents a better adjustment comparing with model 1 (linear analysis).

Table 9 - Simple Regression R^2

Model 1 – Coefficients[a] Linear				
Model	B	Std. Error	t	Sig.
1 (Constant)	−0.11587	0.02041	−5.676	2.37e−08***
Workload	0.21385	0.02408	8.880	< 2e−16***
Residual standard error: 0.9859 on 525 degrees of freedom				

Model 2 – Coefficients[a] – Quadratic				
Model	B	Std. Error	t	Sig.
2 (Constant)	−0.17023	0.02755	−6.180	1.36e−09***
Workload	0.20236	0.02422	8.354	6.96e−16***
Workload * Workload	.07382	0.02535	2.913	0.00375**
Residual standard error: 0.9779 on 524 degrees of freedom				

Source: secondary CNJ data (using R).

Table 10 - Simple Regression Intercept and Coefficients

Model 1 – Coefficients[a] Linear				
Model	B	Std. Error	t	Sig.
1 (Constant)	−0.11587	0.02041	−5.676	2.37e−08***
Workload	0.21385	0.02408	8.880	< 2e−16***
Residual standard error: 0.9859 on 525 degrees of freedom				
Model 2 – Coefficients[a]– Quadratic				
Model	B	Std. Error	t	Sig.
2 (Constant)	−0.17023	0.02755	−6.180	1.36e−09***
Workload	0.20236	0.02422	8.354	6.96e−16***
Workload * Workload	.07382	0.02535	2.913	0.00375**
Residual standard error: 0.9779 on 524 degrees of freedom				

Source: secondary CNJ data (using R).
* p < .05. ** p < .01., *** p < ,001.

The Analysis of Variance in Table 11 compare both models, and It resulted in Degree of Freedom 1 (difference in the number of parameters) with a p-value of less than 0.01. The F limit table test was also executed in R giving the result of 1.324, should be compared with 8.4832), that means that the second model, with Quadratic, is better than the former model, what confirms the H2b. A caveat raised here because the measurement error is not considered, as any OLS, and therefore there is not perfect reliability (Moosbrugger, Schermelleh-engel, Kelava, & Klein, 2009).

Table 11 - ANOVA Table

Model	Df	RSS	Sum of Sq	F	p-value
1– Linear	487	97.439			
2 – Quadratic	486	95.767	1.6716	8.4832	0.003749**

Source: secondary CNJ data (using R).

According to Moosbrugger et al. (2009), the two most essential points in non-linear regressions are measurement error and multicollinearity. But nature measurement error is not used in regressions then just SEM is possible. Additionally, in our study case the reliability of Workload, Cronbach's alpha, Composite Reliability, and AVE are one (1) because it is a single-factor construct. Thus the reliability is not possible to be measured correctly. The multicollinearity between the linear factor and the non-linear factor is a problem because correlation was 0.946 for uncentered (raw data), and 0.549 for centered (scores data), both high correlations. The centralization is useful to avoid non-essential correlation, and the essential-correlation cannot be avoided (Moosbrugger et al., 2009).

Assuming the frailties of regression, we assumed the SEM as a better option, and again SmartPLS was executed with the quadratic term, showing significance (0.009) also confirming the H2b Hypothesis. In summary, the performed analysis of SEM resulted in the confirmation of the hypothesis H1 and the rejection of the other two hypotheses (H2a, H3), and quadratic analysis confirmed the H2b, as shown in Table 12.

Table 12 - Research hypotheses Summary

Hypotheses	Description	Result
H_1	Workforce (Judges, Servers, Commissioners, Outsourcers, and Interns) has a direct positive effect on Court Productivity.	Confirmed
H_{2a}	Workload has a direct and negative effect on Productivity.	Rejected
H_{2b}	Workload has a Quadratic effect on Productivity.	Confirmed
H_3	Workload has a moderating effect decreasing the causal relationship between Workforce and Productivity.	Rejected

Source: secondary CNJ data.

4. Discussions and conclusions

About the examination of different contexts, where Brazil is compared with Slovenia, we suggest that there are benefits in integrating the contexts to the development of a unified theory. However, this integration would not be possible for the current work since it uses secondary data from the National Council of Justice and therefore it would not be possible to have a direct relationship with the cornerstone work. From this observation emerges an opportunity for future works using surveys in both countries, which could better explain some ambiguities emanated by contextual questions, differentiating them from free of context causality.

The information about the weight of the average expenditure in Table 5 is undoubtedly fragile and is a limitation for IPMA analysis, but this analysis gives some opportunities for future approaches like an advanced analysis of weights using experts or other characteristics aside money expenditure.

The present empirical evidence on judiciary productivity is still minimal and not conclusive; this work tried to contribute to a parsimonious model involving constructs which will result in a stronger Performance R^2. Thereby the direct effect of workload, or its moderation, will be out of the model because the results bring no significant influence of this construct.

The present work already points out clues, like the quadratic analysis, for the development of future works that could evaluate antecedents or consequents of workload construct. An additional suggestion for future studies is longitudinal research for Workload, including the quadratic on Productivity. Notwithstanding, assuming that here it is not possible to carry out axiological evaluations regarding labor issues related to workload, we point that psychological impacts, in judges or other profile, could also be evaluated in future works, especially in qualitative research.

About the first hypothesis on the impact of the workforce on productivity, the result showed that it is robust and significant, what contradicts the findings of Dimitrova-Grajzl et al. (2012). The cornerstone work also showed that workload is a direct determinant of performance, and as we tested in hypothesis H2a, it is statistically denied in the present study. However, the organizational and cultural context of the studies may have influenced the results. About the second hypothesis, which deals with the impact of workload in productivity, it is theoretically consistent, and we advanced in the discussion of Dimitrova-Grajzl et al. (2012) about the endogeneity of these constructs. We assumed that the results of our study tried to minimize the endogeneity argument. Notwithstanding, some technics to test endogeneity are opportunities for future works like the Granger causality test; instrumental variable, and another is to threat the CNJ's dataset as longitudinal and (or) as a panel data because there are different years and profiles (State and Federal Courts). We alert that Smart PLS do not have feature for these tests cited, and trade-offs about the software need to be analyzed better in future. We gave a special attention to the workforce construct; it is necessary to consider

the limitation of the present work, which employed only the use of quantitative values, and not taking into account the different levels of expenditure for these profiles of labor in the PLS analysis as we begin with IPMA approach. Despite these limitations, this study presents a steady coefficient of determination (R^2) equivalent to 0.938 and is considerably different from the reference work.

The robust R^2 results are less explained by other constructs, which seems to expand knowledge in the Brazilian context or is biased by CNJ data. It is also apparent that from the legislator's point of view, such as the National Council of Justice, our study provides significant clues to understand the most critical investments in the workforce.

References

Antonakis, J., Bendahan, S., Jacquart, P., Lalive, R. (2010). On making causal claims: A review and recommendations. *The Leadership Quarterly*, 21(6), 1086–1120. https://doi.org/10.1016/j.leaqua.2010.10.010.

Bettis, R., Gambardella, A., Helfat, C., Mitchell, W. (2014). Quantitative empirical analysis in strategic management. *Strategic Management Journal*, 953(April), 949–953. https://doi.org/10.1002/smj.

Capaldo, G., Costantino, N., Pellegrino, R., Rippa, P. (2018). The Role of Risk in Improving Goal Setting in Performance Management Practices within Public Sector: An Explorative Research in Courts Offices in Italy. *International Journal of Public Administration*, 41(12), 986–997. https://doi.org/10.1080/01900692.2017.1317799

CNJ (2016). Justiça em Números 2016 (ano-base 2015). Retrieved November 27,2016, from: http://www.cnj.jus.br/files/conteudo/arquivo/2016/10/b8f46be3dbbff344931a933579915488.pdf;

Cohen, J., Cohen, P., West, S. G., Aiken, L. S. (2003). Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences (Third). Mahwah, NJ: Lawrence Erlbaum Associates.

Dimitrova-Grajzl, V., Grajzl, P., Sustersic, J., Zajc, K. (2012). Court output, judicial staffing, and the demand for court services: Evidence from Slovenian courts of first instance. *International Review of Law and Economics*, 32(1), 19–29. https://doi.org/10.1016/j.irle.2011.12.006.

Falavigna, G., Ippoliti, R., Manello, A., Ramello, G. B. (2015). Judicial productivity, delay and efficiency: A Directional Distance Function (DDF) approach. *European Journal of Operational Research, 240*(2),592-601.http://dx.doi.org/10.1016/j.ejor.2014.07.014.

Fox, J. Weisberg, S. (2011). An {R} Companion to Applied Regression, Second Edition. Thousand Oaks CA: Sage.URL: http://socserv.socsci.mcmaster.ca/jfox/Books/Companion.

Gomes, A. O., Alves, S. T., Silva, J. T. (2018). Effects of investment in information and communication technologies on productivity of courts in Brazil. *Government Information Quarterly, 35*(3), 480-490. https://doi.org/10.1016/j.giq.2018.06.002.

Gomes, A. de O., Guimarães, T. de A. (2013). Desempenho no Judiciário. Conceituação, estado da arte e agenda de pesquisa. *Rev. Adm. Pública*, 47(2), 379–401.

Gomes, A. O., Guimaraes, T. A., Akutsu, L. (2017). Court Caseload Management: The Role of Judges and Administrative Assistants. *RAC - Revista de Administração Contemporânea*, 21(5), 648–665. https://doi.org/http://dx.doi.org/10.1590/1982-7849rac2017160179.

Guimaraes, T. A., Gomes, A. O., Guarido, E. R., Filho (2018). Administration of justice: an emerging research field. *RAUSP Management Journal*, 53(3), 476–482. https://doi.org/10.1108/RAUSP-04-2018-010.

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., Tatham, R. L. (2009). Analise multivariada de dados. (Bookman, Ed.), Bookman (6th ed.). Porto Alegre: Bookman.

Hair, Hult, G. T. M., Ringle, C., Sarstedt, M. (2017). A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM). Thousand Oaks: Sage.

Head, M. L., Holman, L., Lanfear, R., Kahn, A. T., Jennions, M. D. (2015). The Extent and Consequences of P-Hacking in Science. *PLOS Biology*, (March), 1–16. https://doi.org/10.1371/journal.pbio.1002106.

Jaccard, J., Jacoby, J. (2010). Theory construction and model-buildind skills: a practical guide for social scientists. New York, NY: The Guilford Press. Retrieved from www.guilford.com.

Louro, A. C., Santos, W. R., Zanquetto, H., Filho (2017). Productivity antecedents of Brazilian courts of justice: Evidence from Justiça em números. *BAR - Brazilian Administration Review*, 14(4). https://doi.org/10.1590/1807-7692bar2017170032

Mitsopoulos, M., Pelagidis, T. (2007). Does staffing affect the time to dispose cases in Greek courts? *International Review of Law and Economics*, 27(2), 219–244. https://doi.org/10.1016/j.irle.2007.06.001.

Moosbrugger, H., Schermelleh-engel, K., Kelava, A., Klein, A. G. (2009). 6. testing multiple nonlinear effects in structural equation modeling: a comparison of alternative estimation approaches. *In T. Teo M. S. Khine (Eds.), Structural equation modelling in educational research: Concepts and applications*. Rotterdam, The Netherlands: Sense. (pp. 103–136).

Ng, G. Y. (2011). A discipline of judicial governance? *Ultrecht Law Review*, 7(1), 102–116.

Nogueira, J. M. M., Oliveira, K. M. M. de, Vasconcelos, A. P. de, Oliveira, L. G. L. (2012). Estudo exploratório da eficiência dos Tribunais de Justiça estaduais brasileiros usando a Análise Envoltória de Dados (DEA). *Rev. Adm. Pública*, 46(5), 1317–1340. https://doi.org/10.1590/s0034-76122012000500007.

Procopiuck, M. (2018). Information technology and time of judgment in specialized courts: What is the impact of changing from physical to electronic processing? *Government Information Quarterly*. https://doi.org/10.1016/j.giq.2018.03.005.

R Core Team (2018). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Sadek, M. (2004). Poder Judiciário: perspectivas de reforma. Opinião Pública, X, 1–62. Retrieved from: http://www.scielo.br/scielo.php?pid=S0104-62762004000100002&script=sci_arttext;.

Teixeira, J. C., Bigotte, J. F., Repolho, H. M., Antunes, A. P. (2018). Location of courts of justice: The making of the new judiciary map of Portugal. *European Journal of Operational Research*. https://doi.org/10.1016/j.ejor.2018.06.029.