

JUDGES' PERSPECTIVE ON THE LEVEL OF PUNISHMENT

Moshe Bar Niv* & Ran Lachman†‡

Whether or not courts impose an adequate level of punishment, is an important issue in terms of sustaining the social order, maintaining the judicial system's legitimacy, and designing anti-crime policies. To assess the level of sentencing the study surveyed longitudinally, the perspectives of Israeli judges on the issue over a period of three decades. The results show that, consistently, the judges assessed the level of punishment as quite lenient. The results also suggest that no corrective action was taken over the three decades to adjust for the lenient sentencing either by the court system or by the judges themselves, who have the discretion to impose more severe sentences. A regression analysis revealed that court instance and tenure as a judge were related to the judges' assessments of punishment. The practical and theoretical implications of all these results are discussed.

Keywords: level of punishment, judges, sentencing, court systems, tenure as judge, gender, homeostasis

TABLE OF CONTENTS

I. INTRODUCTION	172
II. PUBLIC PERCEPTIONS ON THE LEVEL OF PUNISHMENT.....	176
1. <i>Studies on Public Opinion</i>	177
2. <i>Criticism of Public Opinion Studies on Sentencing</i>	181
III. JUDGES' PERCEPTIONS – FINDINGS	186
1. <i>Judges' Assessment of the Prevalent Punishment Level</i>	187
2. <i>A Longitudinal Comparison</i>	188

* Radzyner Law School, The Interdisciplinary Center, Israel

† Prof. Emeritus, College of Management – Academic studies, Israel

‡ We thank the Research Unit at the School of Business, The College of Management (COMAS, Israel) for financing the data collection. The Research Unit at COMAS and the Interdisciplinary Center participated in supporting this study. We are also thankful to Joseph Ganel and Adi Har-Zion for their assistance. We also thank two EJLS anonymous reviewers for their helpful comments on an earlier version of this paper.

IV. JUDGES' BACKGROUND AND THE ASSESSMENT OF THE LEVEL OF PUNISHMENT	192
1. <i>Judges' Assessments of Punishment by their Court Instance</i>	193
2. <i>Judges' Assessments of Punishment Based on Tenure</i>	193
3. <i>Judges' Assessments of Punishment by Previous Employment</i>	195
4. <i>Judges' Assessments of Punishment by Gender</i>	196
5. <i>Multivariate Analyses</i>	197
V. CONCLUSION	199
APPENDIX	202
1. <i>The Research Method</i>	202
2. <i>Survey Procedure</i>	202
3. <i>The Respondents</i>	203
4. <i>Possible Response Biases</i>	204
5. <i>Measuring the Dependent Variable</i>	205

I. INTRODUCTION

Do courts impose an adequate level of punishment? This is a crucial question for sustaining and legitimizing the criminal system, as well as a central issue in designing anti-crime policies. But, how can the adequacy of the level of punishment be assessed? Unfortunately, there is no acceptable empirical yardstick for the assessment of the general level of sentencing. Most empirical analyses of levels of punishment are based upon surveying public opinion. However, this approach is inadequate and unreliable.¹ In the present paper, we examine the level of punishment from a different and unique point of view – that of the judges themselves. By introducing the judiciary's perspective and by empirically examining judges' assessments of the prevalent level of punishment, this paper contributes to the study and understanding of the level of punishment issued by the courts.

The overall severity level of judicial punishment is a very important issue. Punishment is probably one of the most severe governmental interventions in human fundamental rights. It entails a wide range of sanctions, including

¹ See discussion in section B.2. below.

deprivation of freedom by lengthy incarceration and, in extreme cases, even capital punishment. Thus, the adequacy of sentencing is of essential importance for sustaining the courts' societal legitimacy, for the welfare of society at large, and for the preservation of human rights. The proper sentencing level is also essential for the assessment of court effectiveness and is a key factor for upholding the public's trust in the court system. Therefore, there is a keen interest in studying the overall level of courts' punishment (severity or leniency).

Previous studies have shown that the general public perceives sentencing levels as too lenient. However, while the public's perception of punishment is crucial for the courts' social legitimacy, it does not substitute the professional perspective of the legal branch as adequate feedback on the level of punishment for the courts. Thus, by focusing on public opinion, most current research has failed to shed light on the perspective of those professionals that are actually involved in the criminal punishment process such as judges and lawyers. There are few empirical studies on the perception of the key players in the sentencing process: the judges.² Consequently, very little is known about how judges conceive and perceive the level of punishment that they themselves issue. Thus, exploring the judiciary's perceptions on the issue is of high importance, both from a theoretical and practical standpoint. This study is merely a step towards the exploration of this important issue, and it ought to be followed by further research.

This study examines and analyses the perceptions of presiding Israeli trial court judges regarding the level of punishment that is most prevalent in

² The few surveys examining judicial perceptions with regard to the level of sentencing are primarily 'experimental' in nature. Such an example is a sentencing experiment conducted in Dutch criminal courts, see, Jan W. De Keijser, Peter J. Van Koppen, and Henk Elffers, 'Bridging the gap between judges and the public? A multi-method study' [2007] *Journal of Experimental Criminology* 131-161. Similar experiments usually present the same scenario to a group of laypersons as well, and compare their resulting simulated sentences; see, Andre Kuhn, 'Public and judicial attitudes to punishment in Switzerland' in Julian V. Roberts (ed.), *Changing Attitudes to Punishment: Public Opinion, Crime and Justice* (Routledge 2002), 117. See also, Austin Williams, Thomas Williams 'A Survey of Judges' Responses to Simulated Legal Cases' [1977] 68 *Journal of Criminal Law and Criminology* 307.

Israel.³ The Israeli court system and, in particular, the criminal law are based on the common law system. The Israeli Penal Code and the specific offenses defined in it are based upon the English criminal law.⁴ Hence, the results of this study may help improve the understanding of the level of punishment in other common law jurisdictions as well.⁵

-
- ³ The Israeli judicial system has a three instance hierarchy – magistrate, district and supreme courts. The second instance (district court) serves as a trial court (first instance for matters that are beyond the jurisdiction boundaries of the first instance) and as a court of appeal, over the judgments of the first instance – the magistrate court.
- ⁴ The Criminal Code Ordinance 1936 was enacted during the British mandate and was adopted by Israel upon its independence in 1948. In 1977, the code was replaced by the Israeli Penal Law 5737 - 1977. The new law followed the pattern of the English-Mandatory Statute (in particular regarding the specific offenses). The Penal Law comprises two parts – general and a specific. The general part sets the basic principles of the criminal laws and provides the doctrines relating to criminal behavior. The specific part of the Penal Law details and defines the offenses ranging from infractions to the most serious crimes such as treason and murder. Following the common law pattern, offenses are divided into different categories as a function of the severity of the sanctions, and in particular, to the length of the maximum prison sentence per each crime. Hence, the Penal Code defines three categories of offenses: a) Infractions (petty crimes) – crimes that are subject to a maximum of three-month jail term. Jail terms from three months and up to three years are misdemeanors. Crimes that impose a prison sentence of more than three years or death penalty are defined as felonies. Generally, the Penal Law defines the elements of each particular offense: the criminal behavior (*actus reus*), and the required nature of intent (*mens rea*). The specific part also determines the particular sanctions regarding each offense. In addition to the Penal Law, other laws define numerous particular offenses such as the Securities Laws, Consumer Protection Laws, etc.
- ⁵ The Israeli court system is an adversary one and is based upon the common law courts' principles. See for example, 'As for other nations' practices, looking first to a handful of common law countries whose judicial systems most closely resemble our own, specifically the United Kingdom, Canada, and Israel Mary L. Clark, 'Judges Judging Judicial Candidates: Should Currently Serving Judges Participate in Commissions to Screen and Recommend Article III Candidates Below the Supreme Court Level?' [2009] 114 Penn. St. L. Rev. 49, 59; see also, 'United states and Israel share some features (they are both democracies with a common law tradition', Bruce Peabody, *The Politics of Judicial Independence: Courts, Politics, and the Public* (Johns Hopkins University Press 2011) 208-9. See also, 'courts based on the British derived common law adversarial system practiced in Israel', Angeline Lewis, *Judicial Reconstruction and the*

Typically, Israeli criminal laws prescribe a maximum penalty for defined offences (usually a maximum fine and/or maximum prison term). Only seldom does Israeli law impose mandatory⁶ or minimal⁷ jail penalties. Obviously, most crimes are not subject to the maximum penalty; this creates a legal vacuum that is being filled by judge-made law. Striving to facilitate achievement of the punitive goals, the Israeli Supreme Court has laid down very broad standards, requiring that punishment be based on the principles of 'deterrence, retribution, prevention, and rehabilitation'.⁸ These general principles confer upon the trial courts wide discretion, entailing very limited scope of appellate review.

Throughout the period covered by this study (1989-2010/11), criminal sentencing laws were stable. The general part of the Penal Law, however, was revised in 1995. The revision granted courts limited discretion (in exceptional cases) to alleviate mandatory life-sentence penalties. However, as there are only a handful of such mandatory sanctions, and since the certain offenses and sanctions in the specific part were generally left unchanged, the impact of the revision on the general sentencing policy was limited.⁹

Given the above, this study examines the following questions:

What are the trial judges' perspectives on and assessment of the severity level of current courts sentencing?

What are the judges' views of sentencing given a longitudinal perspective of almost three decades?

Do judges' assessments of punishment relate to their individual

Rule of Law (Brill 2012) 96; Yoav Dotan, *Lawyering for the Rule of Law: Government Lawyers and the Rise of Judicial Power in Israel* (CUP 2013) 18ff.

⁶ The Penal Code imposes mandatory life sentence in relation to murder crimes and genocide crimes.

⁷ The Penal Code mandates a minimum imprisonment for assaulting police officers or obstruction to police officers while they carry out their duties.

⁸ See, CA 3417/99 *Har-Shefi v. State of Israel*, section 9 of justice Tirkel's judgment, <http://versa.cardozo.yu.edu/opinions/har-shefi-v-state-israel>.

⁹ A year or so after our research was completed a substantive reform in sentencing policy has been passed by the Israeli legislature. A structured sentencing regime replaced the previous judicial wide-discretion policy, Penal Law (amendment no. 113), 2012, the amendment became effective as a law on 10.7.2013.

characteristics (such as court instance, tenure as judge, gender, etc.)?

The underlying proposition is that the presiding judges will perceive the punishment they issue as well as that of and their colleagues as adequate. The present study examines this proposition empirically.

The focus is on the judiciary as it plays a pivotal role in the sentencing process. Within statutory constraints, judges are the ones who decide on the actual penalties. Their role in the sentencing process requires the highest expertise and experience regarding sentencing policies and the prevalent levels of punishment. Judges have a comprehensive knowledge of penal law, procedure and sentencing principles, as well as familiarity with the particulars of the criminal cases. They are the ones who evaluate the facts, deliberate the legal issues, and take into account the special circumstances of the case in determining the appropriate punishment in each case. The judges' views on the level of punishment are therefore unique and of utmost importance and relevance.

II. PUBLIC PERCEPTIONS ON THE LEVEL OF PUNISHMENT

Basic constitutional and human rights principles, as well as considerations of public legitimacy of courts,¹⁰ require that the level of criminal sentencing be just and adequate. Overly harsh penalties infringe upon the offenders' rights, and those too lenient penalties impinge on victims' rights and the need to protect society from criminal activities. The task of the courts is to properly set punishments within the sentencing ranges prescribed by the law. In so doing, courts serve the social ends of maintaining an optimal level of deterrence,¹¹ as well as maximizing the welfare of society at large.

¹⁰ On the issues of level of sentencing and public confidence and trust in the judicial system, see: Julian V. Roberts, Mojca, M. Plesničar, 'Sentencing, Legitimacy, and Public Opinion' in Gorazd Meško and Justice Tankebe (eds.), *Trust and Legitimacy in Criminal Justice* 33-51 (Springer 2015), see also, Mike Hough, Ben Bradford, Jonathan Jackson, and Julian V. Roberts, *Attitudes to Sentencing and Trust in Justice* (2013).

¹¹ For a recent model of optimal deterrence, see Steven Shavell, 'A simple model of optimal deterrence and incapacitation', [2015] *International Review of Law and Economics* 13-19. See also, Nuno Garoupa, 'Economic Theory of Criminal Behavior' *Encyclopedia of Criminology and Criminal Justice* (Springer 2014) 1280-1286.

Studying the perceptions of various sections of society regarding the level of punishment has important social implications. For example, a prevalent conception among the public is that the level of punishment, regardless of its real effects, may affect the degree of legal compliance, and consequently, affect the crime rates in society. Thus, a public perception that sentencing has been overly lenient may contribute to a increase in crime and have negative consequences for social welfare.

1. Studies on Public Opinion

The most common approach for assessing the perception of the level of sentencing has been through surveying public opinion.¹² The most prevalent methods used to gauge the public's opinion on sentencing are public opinion polls, media polls, focus groups, etc.¹³ In the last decades, several public opinion surveys have been conducted assessing the prevalent levels of punishment in various countries.¹⁴ The findings suggest that, in many

¹² Another, more 'objective' method for evaluating the severity of sentencing is to compare actual sentences to various statistical parameters. For instance, to compare the average (or median) number of years of the actual sentencing to the 'average' range prescribed by law. The difference (if any) between them can suggest the severity or leniency of the actual punishment and gage skewedness, proportions etc. See, for example: Oren Gazal-Ayal, Ruth Kannai, 'Determination of Starting Sentences in Israel-System and Application' [2010] *Federal Sentencing Reporter* 232-242. However, such calculations, in the context of the adequate level of sentencing, are usually meaningless, since they do not relate to accepted criteria of adequate punishment.

¹³ Arie Freiberg, Karen Gelb, *Penal Populism, Sentencing Councils and Sentencing Policy* (Routledge 2014) 69-70.

¹⁴ See for example: Mike Hough, Ben Bradford, Jonathan Jackson, and Julian V. Roberts 'Attitudes to Sentencing and Trust in Justice: Exploring Trends from the Crime Survey for England and Wales' (LSE Report, 2013); Julian V. Roberts, 'Public Opinion and the Nature of Community Penalties: International Findings' In Julian V. Roberts and Mike Hough (eds.) *Changing Attitudes to Punishment* (Routledge 2002) 33; Austin Lovegrove, 'Sentencing and Public Opinion: An Empirical Study of Punitiveness and Lenience and its Implications for Penal Moderation' (2013) 46 *Australian & New Zealand Journal of Criminology* 200-220. Kimberly N. Varma, Voula Marinos. 'Three Decades of Public Attitudes Research on Crime and Punishment in Canada' (2013) 55 *Canadian Journal of Criminology and Criminal Justice* 549-562. Julian V. Roberts, Mike Hough. 'Exploring public attitudes to

countries, the public views the courts' sentencing as lenient or as highly lenient.¹⁵ Only a small proportion of respondents perceived punishments as severe. For example, Hough and Roberts reported that their survey of the British public regarding the level of punishment showed that about four-fifths of the public in the UK perceived the level of punishment as lenient.¹⁶ Half (51 percent) evaluated the level of punishment as much too lenient. Only a minority (19 percent) estimated the level of punishment as appropriate, and just 3 percent of them perceived it as too severe. In other words, the overwhelming majority of the public considered the punishment to be lenient. The British Annual Survey, which examined the attitudes of the public in relation to the level of sentencing, showed similar results (Table 1 below).

sentencing, factors in England and Wales' in Julian V Roberts (ed.), *Mitigation and Aggravation at Sentencing* (CUP 2011) 168; Mike Hough, Julian V. Roberts. 'Sentencing trends in Britain Public knowledge and public opinion' (1999) 1 *Punishment & Society* 11-26.

¹⁵ Ryan Kornhauser, 'Economic individualism and punitive attitudes: A cross-national analysis' (2015) 17 *Punishment & Society* 27-53; 'One of the leitmotifs of public attitudes to criminal justice is the desire for a harsher response to crime. Most people believe that the justice system is too lenient towards offenders. This perception goes back for many years.' Julian V. Roberts, Michael Hough, *Understanding Public attitude to Criminal Justice* (Open University Press 2005) 13; 'For decades the responses have been the same: most people believe that judges are too lenient towards offenders. This widespread dissatisfaction with the severity of sentencing is probably the most replicated finding in the field', *ibid*, 76.

¹⁶ Michael Hough, Julian Roberts, 'Attitudes to Punishment: Findings from the British Crime Survey' [1998] Home Office Research Study 179, 28, <http://www.icpr.org.uk/media/10372/Attitudes%20to%20punishment,%20hors179.pdf> accessed 15 September 2015.

This attitude remained stable for a long time, see, William Dawes, et. al, 'Attitudes to Guilty Plea Sentence Reductions' (Sentencing Council 2011) Ch. 2 https://docs.google.com/viewer?url=https%3A%2F%2Fwww.sentencingcouncil.org.uk%2Fwp-content%2Fuploads%2FAttitudes_to_Guilty_Plea_Sentence_Reduction_s_web1.pdf accessed 15 September 2015.

Table 1: Attitudes toward sentence severity (2008/09 to 2010/11)¹⁷

	2010/11	2009/10	2008/09
Much too tough	1%	1%	1%
A little too tough	2%	2%	2%
About right	24%	22%	21%
A little too lenient	36%	35%	35%
Much too lenient	38%	40%	41%
Un-weighted N (=100%)	5,596	5,389	5,572

It is also worthwhile to note that this perception of leniency is constant and consistent over three years. Hough and Roberts concluded:

This widespread dissatisfaction with the severity of sentencing is probably the most replicated finding in the field. It appears that whenever (and wherever) the public had been asked the question the majority responded in this way.¹⁸

Similar results of the perceived leniency have been also reported in other common-law countries as well as in some civil law countries,¹⁹ and even in the Nordic countries.²⁰ Typical results of surveys in those countries show that the punishment is perceived as lenient, with only a minority of respondents perceiving the level of punishment as appropriate, and significantly fewer

¹⁷ Michael Hough, Ben Bradford, Jonathan Jackson, Julian V. Roberts 'Attitudes to Sentencing and Trust in Justice'(2013).23, Table 3/6, <https://docs.google.com/viewer?url=http%3A%2F%2Fprints.bbk.ac.uk%2F5195%2F1%2F5195.pdf> accessed 15 September 2015; See also, Jane B. Spratt, Cheryl Marie Webster, and Anthony N. Doob. 'Punishment Severity and Confidence in the Criminal Justice System' (2013) 55 Canadian Journal of Criminology and Criminal Justice 279-292.

¹⁸ Julian V. Roberts, Michael Hough, *Understanding Public Attitudes to Criminal Justice* (Open University Press 2005) 76.

¹⁹ Helmut Kury, Joachim Obergfell-Fuchs, and Ulrich Smartt, 'The Evolution of Public Attitudes to Punishment in Western and Eastern Europe' in Julian V. Roberts, Michael Hough (eds), *Changing Attitudes to Punishment: Public Opinion, Crime and Justice* (2005) 93-114.

²⁰ Balvig Flemming, Helgi Gunnlaugsson, Kristina Jerre, Henrik Tham, and Aarne Kinnunen, 'The Public Sense of Justice in Scandinavia: A Study of Attitudes towards Punishments' (2015) 12 European Journal of Criminology 342-361.

respondents seeing the punishment as severe. In the United States, for example, public opinion surveys indicate that the public wishes the courts would deal more severe punishment than is currently being imposed.²¹

Some studies have indicated that there is a gap between the public's perception of punishment leniency and its perception of a 'proper punishment'. In these studies, respondents were asked to suggest appropriate penalties to hypothetical cases (which were based on real cases already determined by the court). The respondents typically suggested a harsher punishment than that actually imposed by the court. Such results raise a question of the value of public opinion polls for determining the level of punishment.²²

Very few surveys were conducted in Israel on the perceived severity of sentencing. The surveys that have been undertaken indicate that the public's attitude is that punishment of offenders is too lenient. A recent poll, surveying a representative sample of the Jewish population in Israel, showed that the majority (70 percent) of the respondents thought that sentencing had been too lenient. Only 10 percent estimated punishment as too severe.²³ Another survey, which focused on offenses against children, also illustrated that the public regards sentencing for such offenses as being too lenient.²⁴

²¹ Francis T. Cullen, Bonnie S. Fisher and Brandon K. Applegate, 'Public Opinion about Punishment and Corrections' (2000) 58 *Crime and Justice* 1-79, 27: 'the public prefers or, at very least, accepts policies that get tough with the offenders'.

²² Jan W. De Keijser, Peter J. Van Koppen, and Henk Elffers, 'Bridging the Gap between Judges and the Public? A Multi-Method Study' (2007) 3 *Journal of Experimental Criminology* 131-161. See also, Martina Feilzer, 'Exploring Public Knowledge of Sentencing Practices' in Julian V. Roberts (ed.), *Exploring Sentencing Practice in England and Wales* (Palgrave Macmillan 2015) 61.

²³ <http://www.themarket.com/law/1.1707669> accessed 15 September 2015, in Hebrew)

²⁴ Schmidt Hillel Benbenishty Rami, 'Public Attitudes toward Child Maltreatment in Israel' [2011] 33 *Children and Youth Services Review* 1181-1188. For the full report in Hebrew see: https://docs.google.com/viewer?url=http%3A%2F%2Fwww.haruv.org.il%2F_Uploads%2FdbAttachments%2Farticleemail.pdf accessed 15 September 2015.

2. Criticism of Public Opinion Studies on Sentencing

Public opinion surveys have been criticized as an unreliable method of analysis of the level of punishment. While polls may reflect the perception of the public, critics of the approach claim that the general public is incapable of providing a proper assessment of the level of sentencing. They assert that given the complex nature of the sentencing processes, laypeople do not have the capability or skills to accurately grasp legal processes and, hence, cannot provide a proper assessment.²⁵ Several reasons for this have been noted:

A. Incomplete Information

The central criticisms of public opinion polls as providing poor reflection on punishment levels are twofold:

- 1) Legal knowledge – The public lacks the knowledge and (legal) understanding required for a relevant and educated assessment of the case;
- 2) Distorted knowledge – Incomplete information and at times misinformation is common among the public.

According to critics, public opinion surveys are not an appropriate proxy for gauging the severity of existing punishment as the public does not have the knowledge and legal understanding required for properly evaluating the level of punishment. Furthermore, the public does not have access to the particular details of the cases that are of the essence in determining the sentences.²⁶ Criticism of the use of public opinion surveys has intensified in the recent years making 'a very strong case against the validity of survey measurements of public opinion on criminal justice'.²⁷ The current view in criminology literature is that public opinion may be relied upon only when the public is provided with sufficient information and is reasonably able to

²⁵ See, Julian V. Roberts, Mojca M. Plesničar 'Sentencing, Legitimacy, and Public Opinion' [2015] *Trust and Legitimacy in Criminal Justice* 33, 43 and reference.

²⁶ See, Neil Hutton, 'Beyond Populist Punitiveness' (2005) 7 *Punishment and Society* 243.

²⁷ Jan W. de Keijser, Henk Elffers 'Cross-jurisdictional differences in punitive public attitudes?' (2009) 15 *European Journal on Criminal Policy and Research* 47-62, 49.

deliberate the questions.²⁸ But in most surveys, this is not the case. Hence, serious doubts have been raised about the validity of public opinion surveys as a dependable or even as a proper means for assessing the levels of punishment in a society. Yet, such polls remain the most common tool for assessing levels of punishment and they still have a significant impact on politician, administrators and others. This may affect future levels of punishment and establish an 'unjustified and unhealthy level of dominance in the contemporary political sphere'.²⁹

As for the problem of distorted and incomplete information, critics have argued that the opinions and perceptions of the public are derived from anecdotal, sporadic, partial, distorted, and/or erroneous information. Hence, the perspective of the public is inaccurate and biased. For the most part, the information available to the public tends to be derived from the mass media,³⁰ which typically focuses on the exceptional, unusual, otherwise sensational, or even inflammatory cases rather than the usual, ordinary court rulings. Consequently, people who are exposed to such media are more likely to consider punishments as too lenient and tend to advocate harsher punitive

²⁸ David A. Green, 'Public Opinion versus Public Judgment about Crime Correcting the 'Comedy of Errors'' (2006) 46 *British Journal of Criminology* 131-154. Julian V. Roberts, "'Community views of sentencing: Intuitive and principled responses to offending' in Michael Tonry (ed.), *Punishment Futures. Studies in Penal Theory and Philosophy* (OUP 2011) ; see also, Julian V. Roberts, 'Clarifying the Significance of Public Opinion for Sentencing Policy and Practice' in Ryberg, Jesper, and Julian V. Roberts (eds.) *Popular Punishment: On the Normative Significance of Public Opinion* (OUP 2014) 228. Jan W. de Keijser, 'Penal Theory and Popular opinion: The Deficiencies of Direct Engagement' in Ryberg, Jesper, and Julian V. Roberts (eds), *Popular Punishment: on the normative significance of public opinion* (OUP 2014) 101, 104ff.

²⁹ Liz Turner, 'Penal Populism, Deliberative Methods, and the Production of 'Public Opinion' on Crime and Punishment' (2014) 23 *The Good Society* 87-102. See also, Matthew Smith, 'The Chicken or the Egg: What Shapes Public Opinion on Punishment?' [2013] *Chicago Policy Review* <http://chicagopolicyreview.org/2013/10/08/the-chicken-or-the-egg-what-shapes-public-opinion-on-punishment/> accessed March 11 2017. It argues that public opinion has caused stricter punishment.

³⁰ On the implication of the media upon respondents in public polls see, Julian V. Roberts et al, *Penal Populism and Public Opinion: Lessons from Five Countries* (OUP 2002) ch. 5.

measures compared to people with less media exposure.³¹ Furthermore, less knowledgeable people rely more heavily on the media's interpretations as a source of information.³²

Several scholars have referred to this issue as the 'punitive gap' - the gap between the public's attitudes on sentencing and the actual sentencing.³³ This gap is attributed in part to the lack of accurate and valid information. In general, the public lacks knowledge and understanding of criminal law and sentencing policy.³⁴ Consequently, the public is incapable of analyzing what information is relevant and form an educated opinion. Various studies have shown that providing laypeople with relevant case information, even as trivial as the range of the statutory sentencing penalties or sentencing alternatives, have changed the respondents' perceptions on severity.³⁵ Furthermore, studies have shown that additional information affects respondents' views on punishment levels.³⁶ Such information make respondents more

³¹ Franklin D. Gilliam, Shanto Iyengar, 'Prime Suspects: The Influence of Local Television News on the Viewing Public' (2000) 44 *Am. J. Pol. Sci.* 560, 563-64.

³² Justin T Pickett, Christina Mancini, Daniel P. Mears, and Marc Gertz, 'Public (Mis) Understanding of Crime Policy the Effects of Criminal Justice Experience and Media Reliance' [2015] *Criminal Justice Policy Review* 500-522. Kenneth Dowler, 'Media Consumption and Public Attitudes toward Crime and Justice: The Relationship between Fear of Crime, Punitive Attitudes, and Perceived Police Effectiveness' (2003) 10 *J. Crim. Justice Popular Culture* 109; see also Julian V. Roberts & Anthony N. Doob, 'News Media Influences on Public Views of Sentencing' (1990) 14 *L. & Hum. Behav.* 451, 456.

³³ Jan W. de Keijser, Henk Elffers, 'Cross-jurisdictional differences in punitive public attitudes?' (2009) 15 *European Journal on Criminal Policy and Research* 47-62, 48ff and references there.

³⁴ At an early stage of the empirical studies of the perceptions of punishment, scholars noticed that there was a substantial gap between the actual levels of punishment and the public's perception of the sentencing; See, Arnold M. Rose, Arthur E. Prell, 'Does the punishment fit the crime? A Study in Social Valuation' [1955] *American Journal of Sociology* 247-259, 248.

³⁵ Julian V. Roberts et al, *Penal Populism and Public Opinion: Lessons from Five Countries* (2005) 29-31 and references to studies there.

³⁶ David Indermaur, Lynne Roberts, Caroline Spiranovic, Geraldine Mackenzie, and Karen Gelb, 'A matter of judgement: The effect of information and deliberation on public attitudes to punishment' (2012) 14 *Punishment & Society* 147-165.

knowledgeable and decrease the 'punitive-gap' in their perceptions of pensiveness.³⁷

To recapitulate, assessments by the public of the level of punishment are, in fact, imprecise impressions based on partial, selective, and even inaccurate information. Some studies show that the public has no meaningful knowledge on the actual level of punishment.³⁸ Scholarly criticism has asserted that public opinions on this issue are unreliable and cannot reflect whether punishment is actually³⁹ too lenient or harsh.⁴⁰

B. The Complexity of Criminal Sentencing

Criminal sentencing is a process of balancing. On one hand, there are elements specific to each case and each offender and, on the other hand, there are public interest and the structure of the statutes. This creates inherent difficulties in having non-professionals assess the level of sentencing. The regular structure of criminal statutes prescribes the maximum and the minimum penalty, and sometimes prescribes no substantial punishment. Judges determine a particular punishment following judicial guidelines and personal discretion. These particular punishments result in a disparity among penalties imposed upon offenders for the breach of the same statutory

³⁷ Arie Freiberg, Karen Gelb, *Penal Populism, Sentencing Councils and Sentencing Policy* (Routledge 2014) 74-76 and references there.

³⁸ See for example, Michael Hough, Julian V. Roberts, 'Sentencing Trends in Britain: Public Knowledge and Public Opinion' (1999) 1 *Punishment and Society* 11-26, 20. See also, Trevor Sanders, Julian V. Roberts, 'Public Attitudes Toward Conditional Sentencing: Results of a National Survey' (2000) 32 *Canadian Journal of Behavioral Science* 199-207.

³⁹ Jesper Ryberg, Julian V. Roberts, *Popular Punishment: On the Normative Significance of Public Opinion* (OUP 2014) 5.

⁴⁰ This, however, does not imply that public opinion surveys are irrelevant. Understanding public perception of punishment is of importance because it may actually influence the behavior of the stakeholders such as potential offenders, lawyers etc. Hence, notwithstanding the real level of punishment, a perception of harsher punishment may deter people from performing socially negative activities, while a public perception of leniency may fail in achieving sufficient deterrence and consequently result in higher rates of criminal activities. In addition, other ramifications may result from the publics' perceptions of punishment such as people's attitudes to fairness and justice.

criminal rule. The inherent variability of punishments is often viewed by non-professionals as inconsistency. Non-professionals tend to view as lenient penalties that are inconsistent with the more severe ones in the range, and this reinforces their perceptions on severity of punishment. A layman's perception, therefore, is of questionable value in assessing overall levels of punishment. In contrast, professionally qualified evaluators, such as judges or lawyers, assess such diversity differently and make a much better (or, at least a less distorted) estimation of the appropriate level of punishment. As indicated earlier, there are very few empirical studies of views expressed by judges or representatives of the legal professions (such as lawyers). Consequently, the primary source of evaluation of the level of punishment is currently based upon the not very reliable grounds of opinion surveys of attitudes of the general public, or segments of it.

C. Limited Accessibility for Research

The focus on public opinion rather than on judges' perceptions may be a result of the practical and legal difficulties of conducting surveys among judges.⁴¹ This is particularly the case when the subject matter of the research may, directly or indirectly, be seen as a criticism of the judicial system. Furthermore, judges are frequently overloaded with work,⁴² and their predisposition to participate in independent academic studies is low. Therefore, it is not surprising to find only a handful of systematic empirical studies of the judges' views regarding punishment.⁴³

⁴¹ To run our surveys, we had to obtain permission from the Chief Justice of the Israeli Supreme Court, who is considered the head of the judicial system.

⁴² See for example, Shay Lavie, 'Appellate Courts and Caseload Pressure' (2016) 27 *Stan. L. & Policy Rev.* 57; Andrew Tickell, 'More 'efficient' justice at the European Court of Human Rights: but at whose expense?' [2015] *Public Law*, 206; Roger J. Miner, 'Dealing with the Appellate Caseload Crisis: The Report of the Federal Courts Study Committee Revisited' (2013) 57 *N.Y.L. Sch. L. Rev.* 517; Michael C. Gizzi, 'Examining the Crisis of Volume in the U.S. Courts of Appeals' (1993) 77 *Judicature* 96.

⁴³ Searching the legal data bases show that, indeed, a small number of surveys of judges' perceptions and attitudes were published, but none dealt with the subject of severity of the level of punishment.

III. JUDGES' PERCEPTIONS – FINDINGS

Our study examined the perceptions on punishment by presiding judges in Israel over a period of nearly 30 years.⁴⁴ Three consecutive surveys were conducted, 10 years apart, among all the first and second-instance judges presiding at the time of the surveys: in 1989, 1999 and 2010.⁴⁵ In the three surveys, the judges were asked to assess the severity level of the punishment prevalent at the time of each survey.⁴⁶ The very same question was used in all three surveys in order to allow a direct comparison. To avoid a self-serving response bias, the judges were not asked to assess their own sentencing but to assess the level of sentencing severity prevailing in Israel at the time.

The question the judges were asked in all three surveys was: '*Given your familiarity with the Israeli courts system, to what extent do the sentences imposed by the courts reflect a lenient approach?*'.⁴⁷ The response range to this question was: 1. *not lenient at all*; 2. *lenient to a small extent*; 3. *partly lenient*; 4. *considerably lenient*; 5. *lenient to a large extent*; 6. *very lenient*.⁴⁸

⁴⁴ Even though the data was collected in Israel, the results may have ramifications for a wider scope of common law judicial systems, of which the Israeli system is part.

⁴⁵ For a detailed description of the research methods and surveys conducted, see Appendix A.

⁴⁶ We acknowledge the possibility of a 'pro-system' bias in judges' assessments of punishment, as their responses may 'tarnish' the reputation of the judicial system and of their colleagues, and perhaps indirectly also implicate themselves. Therefore, the answers of the judges may be biased. If such bias exists, the result will be that the assessment of the punishment level would be that it is more 'appropriate' than it actually is. Consequently, the answers here are perhaps even more 'restrained' than the real opinion is.

⁴⁷ What is meant by 'lenient' was not pre-defined for the judges, but was left for the assessment of the responding judge. For a full discussion of this point, see Appendix A: Research Methods.

⁴⁸ An exploratory study conducted prior to the survey included the categories of '*adequate*' and '*too harsh*'. Yet it was found that no one chose these responses, and they remained 'empty categories'. Hence, they were omitted from the survey. For further details regarding the choice of this kind of response range, see Appendix A: Research Methods.

I. Judges' Assessment of the Prevalent Punishment Level

In the presentation of our results, we will first analyze the findings of the most current survey, conducted in 2010, and then compare them to those attained a decade and two decades earlier.

Our assertion was that since judges are the ones who issue the sentences and decide on the punishment they consider most appropriate, they would assess the level of punishment as fitting and appropriate and not as lenient. The findings did not support this assertion (See Chart 1 below). They showed that most judges did not consider the level of punishment to be appropriate. The majority of them (70 percent) considered punishment to reflect a lenient or at least partly lenient approach. Only very few judges (8 percent) considered the level of punishment to be not lenient at all.

Over a third of the judges (36 percent) thought that the punishments handed out were considerably to very lenient, of whom nine percent evaluated the punishments as lenient to a large and a very large extent. Another third (34 percent) said the existing punishment was partly lenient and 23 percent said it was lenient to a small extent (Chart 1 below).

In other words, even though the judges are the ones who issue the sentences and decide on the punishment, a large majority of them thought that the courts' sentences reflected some level of leniency. We find these results quite surprising. It is commonly assumed that judges are using their best judgment to decide and determine the punishment they think is the most appropriate one given the merits of the specific case. Yet, the results here suggested that the judges did not have confidence in the judgment of their colleagues; rather, they regarded the colleagues' sentencing as unfitting and lenient.

These results also raise an interesting issue: if most judges think that the overall level of punishment tends toward leniency, why do they not adjust it by issuing more adequate i.e., less lenient punishments? After all, it is up to them to decide and in their power to implement it.

Before we engage in finding possible interpretations for these results, it ought to be examined if the results obtained here are not an aberration. Perhaps the severity level of punishment found in 2010 reflected a temporary one-time deviation from the appropriate level. In other words, the level of punishment may follow a pattern of 'dynamic equilibrium'. Like any other 'open system',

the level of punishment is not unwavering or consistently stable, but it fluctuates or oscillates over time around a certain average level: i.e., governed by homeostasis.⁴⁹ In other words, when the level of punishment becomes more lenient than expected, judges may adjust by imposing more severe punishments to 'compensate' and correct for it. This 'correction' or adjustment may oscillate to a level of overly severe punishments, which in turn will lead to a further 'correction' toward higher leniency, to adjust accordingly, and so on. To wit, over time, the level of punishment may constantly fluctuate between too lenient or too severe, depicting a 'dynamic equilibrium' around the adequate punishment. Hence, it may well be that the judges' evaluation of the level of punishment in the 2010 survey was a 'snapshot' at a single point in time and not a dynamic picture of the levels of punishment that oscillate toward leniency as a reaction to a prior period when the perception of severity was much higher. To test this interpretation, a longitudinal perspective will be taken.

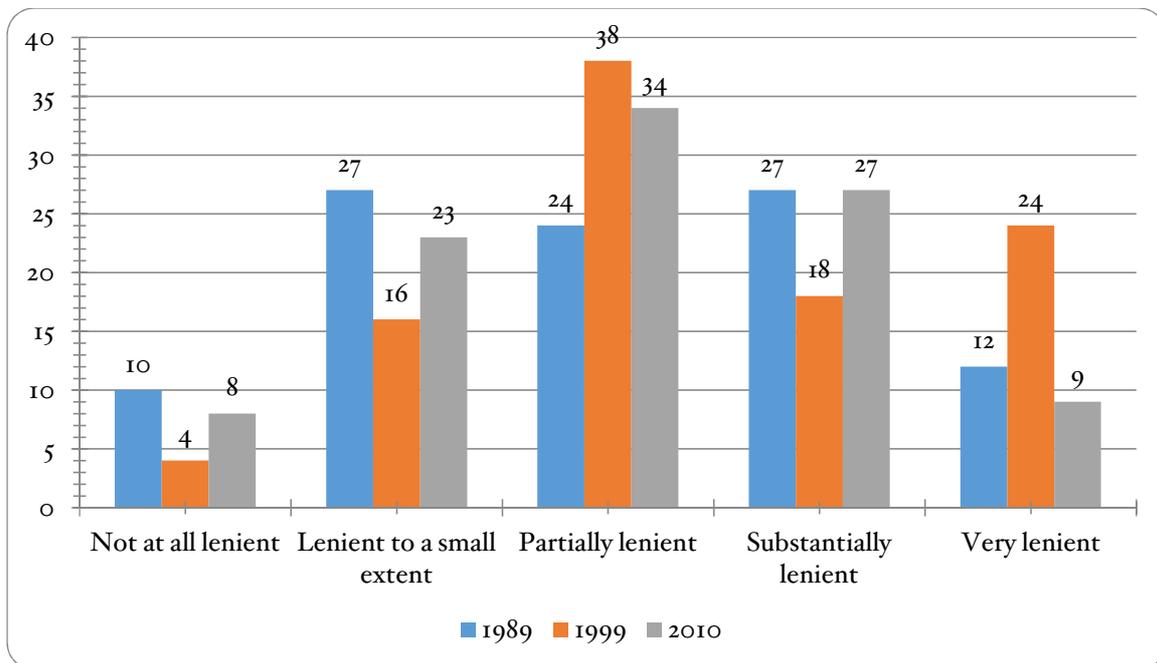
2. A Longitudinal Comparison

In order to test the above interpretation, we examined the perceived level of punishment over an extended time period. As noted above, judges' assessments of the level of punishment were measured not only in 2010, but also one decade earlier (1999) and two decades (1989) earlier. Given the conceptual framework of 'dynamic equilibrium' (or homeostasis), the perception of punishment as lenient may be a 'response' to a perceived severe level of punishment in the preceding decade (1999) and a 'correction' for it. If this is so, the level of punishment as perceived by judges in 1999 can be expected to be more severe, or in terms of our study, as 'not lenient at all'. The same pattern may be assumed for the judges' assessments given two decades earlier, in 1989.⁵⁰ The responses given at the three points in time are compared below (Chart 1).

⁴⁹ 'Homeostasis' is a mechanism inherent within open systems to assure that deviations beyond a given range from their desired course of affairs are self-corrected, thus assuring the systems' survival and sustainability.

⁵⁰ Obviously, there is the question of what is the proper time-cycle that captures the oscillation cycle. We refer to this issue later.

Chart 1: Judges' assessment of level of punishment over three decades (in %)



Looking at the data of the 1999 survey first, the results do not appear to support the 'self-correction' (homeostasis) assertion. In the 1999 survey, the assessments of punishment level, in general, were found to be no different than those in 2010 (Chart 1). The observed differences between them were not statistically significant, suggesting that the two distributions of responses of 2010 and of 1999 – are practically the same.⁵¹ To wit, the judges' perceptions of the punishment level in 1999 pointed towards leniency with the same strength as in 2010.

While these differences in distributions were not statistically different, in 1999, a somewhat larger majority of the judges (80 percent vs. 70 percent in 2010) said that the punishment prevalent at the time reflected leniency to at least some extent. More specifically, in 1999, a quarter (24 percent) of the judges said that the punishments were very lenient, as opposed to only 9 percent who suggested that in 2010.⁵² Similarly in 1999, 42 percent of the

⁵¹ $\chi^2 = 6.51$, d.f. = 8, ($p > 0.10$) n.s.

⁵² The value of statistical significance is affected by the number of observations at hand. Since the number of judges here is rather small this may affect the calculated significance of the differences. Perhaps with a larger N the differences could turn out to be significant. Hence, we decided to describe the differences.

judges assessed the punishment to be lenient 'to a large extent' or 'very lenient', whereas in 2010, 36 percent of the judges thought so. At the other end of the scale, in 1999 only 20 percent of the judges thought punishments were 'not at all lenient' or lenient 'to a small extent only', as opposed to 31 percent of the judges who felt this way in 2010.

These results suggest that the perceived leniency of punishment in 2010 was not a reaction to the perceived level of punishment during the preceding decade. The assertion, thus, that self-correction mechanisms are at work here (homeostasis) is not supported. However, since we have no precise knowledge of the timespan of the oscillating cycle (if one indeed exists), homeostasis cannot be ruled out entirely. It can still be argued that the correction or adjustment cycle takes less (or perhaps more) than a decade, or that we have 'missed' its picks with our surveys, and hence it was not captured by the surveys conducted 10 years apart. We have no data to clearly rule out such an argument.

It can, however, be assumed that adjustment periods, or oscillation cycles, from lenient to harsh punishment and back are not 'instantaneous' and may take several years to occur. Similarly, it may take judges several years of experiencing 'inappropriate' levels of punishment to react accordingly and adjust their own sentencing (if at all). Thus, we suggest that comparing the judges' responses at two points in time may miss or fail to reflect the 'full' cycle of oscillation from lenient punishment. Therefore, we added to the analysis the responses of the judges to this question collected in the 1989 survey. Now, the pattern of perceived level of punishment can be examined over three points in time.

The 1989 results remained basically the same – the distribution of responses was not statistically significant from those of two preceding surveys.⁵³ The majority of the judges in the 1989 survey thought that punishment was at least 'partly lenient', and a minority said it was only slightly lenient or not at all lenient (Chart 1). More specifically, in the 1989 survey, 63 percent of the judges believed that punishments were at least partly lenient, and among this number, 27 percent said the punishment was considerably lenient, and 12 percent felt it was lenient to a large extent or very lenient. Over a third

⁵³ $\chi^2 = 6.51$, d.f. = 8, $p > 0.10$; N.S.

assessed the punishment as lenient only to a small extent (27 percent) or not lenient at all (10 percent).

Thus, adding the data of 1989 to that of the other surveys shows that the distributions of judges' assessments over the three decades were not statistically different from each other. They all similarly suggested that punishment was persistently perceived not to be adequate but to be quite lenient. Also, these results did not support the assertion regarding homeostasis. There appears to be no self-correction (no homeostasis) mechanism, and the lenient approach in punishment prevailed over the rather long time period of three decades.

However, as suggested earlier, one can still not rule out the argument that the cycle of fluctuations, if they existed, did not fully correspond with our surveys. Since the timespan of such a proposed cycle is unknown, there is the possibility that it occurs within each decade and the three points of measurement here actually tapped it at the very same status: namely, a peak of leniency. While we think, it is rather unlikely that all three surveys happened to 'miss' the cycle, our data is not sufficient for ruling out such an event. From the statistical point of view, however, all three distributions, while not identical, were not significantly different.

The presentation of the responses in Chart 1, even though not statistically significant, can be interpreted as reflecting some level of oscillations of decreasing and increasing leniency over the three periods of time. For example, if one examines the responses judges gave in the 1989 survey, 37 percent of them said that punishment was not at all (or almost not) lenient; in 1999 their proportion was reduced to 20 percent, and in 2010, it increased again to 31 percent. Perhaps more pronounced were the differences in proportion of judges who felt the punishment was very lenient or lenient to a large extent: 12 percent in 1989 growing to 24 percent in 1999 and dropping back to 9 percent in 2010. In other words, several judges felt that the level of punishment was less lenient in 1989, became more lenient in 1999 and less lenient again in 2010. Thus, the mechanisms of homeostasis or self-adjustment arguable does exist, correcting the 1989 deviations towards somewhat higher leniency in the level of punishment in 1999, and oscillating it back towards the less lenient level in 2010. Yet, as these observed

differences in responses are not statistically significant⁵⁴ the assertion about a self-correcting mechanism that operates over time in punishments handed out is not supported here.

Given all of the above, we adopt the position that the judges' perception of the level of punishment did not significantly change over the three decades, and that they had been assessing it as lenient all along. We find this consistency in judges' perceptions over such an extended period quite surprising.

IV. JUDGES' BACKGROUND AND THE ASSESSMENT OF THE LEVEL OF PUNISHMENT

The analyses above indicate that the judges' assessments of the levels of punishment were not uniform and not all of them evaluated them in the same manner. Do these different views relate to some professional background factors such as their years of experience, the court they sit on, their employment prior to their appointment to the bench, or perhaps personal factors such as their gender?

Courts,⁵⁵ as well as scholars,⁵⁶ have recognized that a judge's background may affect their decisions. This is particularly true in cases where judges have a wide discretion in sentencing.⁵⁷ We examined some of these possible

⁵⁴ The commonly acceptable probability of error is 0.05% or less, and the probability of error found here is higher than that. Furthermore, when the population at hand is rather small (as the numbers of judges here), a less strict criterion of $p < 0.10\%$ may also be used, but the results reported above ($p > 0.10$) did not meet this lax criterion either.

⁵⁵ As Judge Richard A. Posner recognized: '[T]he exercise of discretion is shaped by a judge's values and intuitions, which in turn are shaped by the judge's background and experiences', *Tyson v. Trigg* 50 F.3d 436, 439 (7th Cir.1995).

⁵⁶ See for example, Paula M. Kautt, Cassia C. Spohn, 'Assessing Blameworthiness and Assigning Punishment', in Edward R. Maguire, David E. Duffee (eds), *Criminal Justice Theory: Explaining the Nature and Behavior of Criminal Justice* (Routledge 2015) 220-221.

⁵⁷ See for example, 'Although a sentencing judge is bound to make the findings and consider the relevant factors as required by the sentencing law, the manner in which a judge performs these duties may be guided by that judge's background, experiences, and moral values', *State v. Brown*, Court of Appeals of Ohio, 2004 WL 764589.

relationships by analyzing the data collected in the most recent survey of 2011, where the number of respondents was the greatest (86 judges).⁵⁸

1. Judges' Assessments of Punishment by their Court Instance

The assessment of the level of punishment may differ by court instance. The first instance (Court-of-Peace) in Israel generally adjudicates the crimes that entail imprisonment of up to a maximum of seven years, while the second instance (District Court) has jurisdiction on felonies that are subject to maximum imprisonment exceeding seven years. The gravity of the crimes adjudicated by each instance may reflect in a different perspective on punishment. Moreover, Israeli second instance courts have a twofold jurisdiction in criminal proceedings: they serve as trial courts within the scope of their vested jurisdiction and also as appellate courts that review first instance judgments. Hence, they have the formal authority and practical experience to evaluate the adequacy of the decisions and sentences of the lower instance courts. This special jurisdictional structure may affect the perception of the level of magistrate (First Instance) judges.

In order to examine such possible effect, the judges were divided as first and second instance, and the distributions of their assessments were compared. The results showed the distributions of responses of judges of the two instances were not different from each other or statistically significant.⁵⁹ Thus, no difference in perceptions of punishment severity exists between first and second instance judges.

2. Judges' Assessments of Punishment Based on Tenure

We further tested if the tenure of a judge could explain difference in perception of leniency. We expected that the more experienced the judge (as measured by years as a judge) the more he or she would assess punishments as lenient. Not only do years on the bench provide a wider perspective on punishment, longer experience allows for a different perspective on recidivism. To test this, years spent as a judge were correlated with the judges' assessment of level of punishment. The results show the two were correlated

⁵⁸ See Appendix A: Research methods, for details.

⁵⁹ Differences are not significant: $\chi^2=3.4$, d.f.=4; $p=0.46$, n.s.

($r_p = 0.25$) with each other.⁶⁰ The proposition is therefore supported: the more experience judges have, the more they tend to assess punishment as lenient. However, the correlation found here is not a strong one, suggesting that judges' experience is only weakly related to their assessments of punishment.

However, the correlation coefficient used here (Pearson's coefficient) taps linear relationships only. But, the relationships between years of tenure and assessment of punishment may be non-linear in nature. In fact, the Pearson coefficient indicates whether relationships between two variables were linear or not. Thus, if the relationships between years of experience and views on punishment are non-linear, the coefficient may be weak or indicate no relationship altogether. To examine if this is the case here, the relationship between punishment assessment and tenure were also examined using cross-tabulation of the two variables. In order to accomplish this, the variable of years of tenure as a judge was clustered into three categories⁶¹:

a. short experience (0-7 years); b. intermediate experience (8-14 years); c. long experience (over 14 years). The categories of 'very high' and 'high' extent of leniency assessment were also collapsed together.⁶² The distributions were then cross-tabulated (Table 2).

Table 2: Assessment of punishment by tenure as judges (in %)

Tenure over 14	Tenure 8-14	Tenure 0-7	Lenient
30.4	30	30.8	To a small extent
21.7	50	42.3	Partially
47.8	20	26.9	Substantially
(N=23) 100%	(N=20) 100%	(N=26) 100%	Total N=69

The results (Table 2) show that the relationships between tenure and assessment of leniency are, indeed, not quite linear. While there were

⁶⁰ Pearson correlation coefficient (r_p) was calculated here: $r_p = 0.25$, $p < 0.04$ (N=69).

⁶¹ Given the relatively small number of judges in the sample (see appendix A) length of tenure had to be collapsed into no more than three categories. To do that the tenure frequency distribution was divided into three almost equal categories with about a third of the respondents in each (26, 20, 23 respectively), yielding the above grouping.

⁶² These categories were collapsed together to avoid empty cells or cells with very small N in the cross-tabulation.

practically no differences in perceived leniency between judges of short and medium tenure, judges with long tenure perceived punishment as much more lenient.⁶³ Nearly a half (48 percent) of the more experienced judges (14 years and more) assessed punishment as at least substantially lenient; about a third (30 percent) perceived it as lenient to a small extent (none said it was not lenient at all). In contrast, the shorter tenured judges (0-7 and 8-13 years' tenure), 20 percent and 27 percent, respectively assessed punishment as substantially lenient and 42 percent and 50 percent assessed punishment as partially lenient. In addition, it is interesting to see that the assessments of punishment as not lenient (or to a small extent only) were unrelated to years of experience: slightly less than a third of the judges, regardless of experience, assessed punishment as not lenient (30, 30, and 31 percent). The distribution of responses in the other categories (partly, considerably, and very lenient) did vary by experience. Thus, the results suggested that experience as judge and punishment assessment, are related in a slightly non-linear relationship.⁶⁴

It can, therefore, be concluded that the hypothesis that tenure is related to punishment assessment of judges, is supported. The results suggest, however, that experience had a differential effect on judges' evaluations: it made no difference for judges who assessed punishment as not lenient, but many years of experience did make a difference for the perception of punishment as considerably or very lenient.

3. Judges' Assessments of Punishment by Previous Employment

It has been proposed that judges who prior to their nomination to the bench were employed in the Ministry of Justice (e.g., Public Prosecution Office) would assess the level of punishment as more lenient than judges who were previously employed in the private sector.⁶⁵ Of interest here were judges who,

⁶³ $\chi^2 = 14.9$, d.f.=8; Sig=0.06. N=69. The significance test here shows that the differences observed here are not statistically significant at the $p=0.05$ level. However, given the small numbers of judges, one can adopt here the less-conservative approach, where significance can be accepted at the $p<0.10$ level. Hence, we consider these differences as significant.

⁶⁴ This may explain the low correlation coefficient observed above.

⁶⁵ *Tyson v. Trigg*, 50 F.3d 436, 439 (7th. Cir.1995), Judge Posner, 'Former prosecutors may have a different bent from former defense lawyers, former lawyers for tort plaintiffs a different bent from former lawyers for insurance companies'.

before being nominated to the bench, served as defence lawyers as distinguished from those who served as prosecutors. The argument is that judges who served as prosecutors will tend to assess the level of punishment as more lenient than those who served as defence lawyers.

The best proxy we had for previous employment was previous employment in the private sector (e.g., private law firms) or previous employment in the government sector (e.g., Public Prosecution Office, Ministry of Justice). We adopted these two categories in order to test the relationship between previous employment and perceived punishment. The distributions of punishment assessments within each employment group were compared.⁶⁶

The results, however, did not support this proposition: no significant differences were found between the responses based on previous employment.⁶⁷

4. *Judges' Assessments of Punishment by Gender*

Does the perception of punishment relate to the gender of the judge? Previous research provided mixed results on this issue.⁶⁸ Some scholars viewed female judges as more 'liberal,' and as more lenient in punishment.⁶⁹ Given the ambiguous results of various studies, we examined in our data whether gender was related or not to the judges' perceptions of severity of punishment.⁷⁰ The results suggested that there were no statistically

⁶⁶ The original response range of six possible responses of this distribution had to be clustered into three main categories to avoid categories with very small numbers of responses or 'empty' ones. 'Partly' and 'considerable extent' were joined into 'partly' and 'a large' and 'very large extents' were joined into 'large extent'. However, there was no significant difference (χ^2 test) between these distributions when the full distribution was examined.

⁶⁷ $\chi^2 = 1.65$, d.f.=4, $p = 0.80$ n.s.; $N = 77$

⁶⁸ See for example, Shanna R. Van Slyke & William D. Bales, 'Gender Dynamics in the Sentencing of White-Collar Offenders' (2013) 26 *Criminal Justice Studies* 168-196; but, see to the contrary, Barbara Palmer, 'Women in the American Judiciary: Their Influence and Impact' (2001) 23 *Women & Politics* 89-99.

⁶⁹ Joanne Belknap, *The Invisible Woman: Gender, Crime, and Justice* (Cengage Learning 2014) 566.

⁷⁰ To examine it, judges were grouped by gender and the distributions of respondents' assessments of the level of punishment were compared.

significant differences between female and male judges in this matter.⁷¹ Both male and female judges similarly assessed the prevalent level of punishment severity.

5. *Multivariate Analyses*

So far, we have examined the relationships between judges' assessment of punishment and some background factors. However, in reality, the effect of each of the background factors on the judges' evaluations is not isolated from the possible effects of the others. To wit, when combined, the effect of each of these factors on the assessment may overlap or be different than their singular effect. Hence, the joint effect of these factors on the assessment and their relative weights ought to be analyzed as well. A multivariate analysis was, therefore, required. This was done by regressing the judges' assessments of punishment on the three major background factors.⁷²

The regression analysis suggested that the tenure as judge and the instance in which they presided were predictors of the variance in punishment assessments (Table 3). Gender was found to have no significant independent effect. These two variables accounted for nine percent of the adjusted variance in severity assessment, a rather low explained variance.

Table 3: Regression of judges' assessment of punishment on background factors

Collinearity Statistics		Sig	T	Standardized Coefficients	Unstandardized Coefficients		Model r ^a
VIF	Tolerance			Beta	Std. Error	B	
		.000	5.340		.465	2.486	(Constant) ^b
1.329	.752	.004	3.001	.40	.016	.049	Years as judge
1.034	.967	.532	.629	.074	.243	.153	Gender
1.293	.773	.032	2.198	.29	.263	.578	District Court

a) Dependent Variable: Sentences reflect a lenient approach

b) Predictors: (Constant), Gender, First/Second Instance, Tenure (Years) as judge
 $R^2 = 0.13$, $Adj. R^2 = 0.09$

⁷¹ $\chi^2 = 3.19$, d.f. = 4, $p = 0.53$, n.s.

⁷² Given the small number of judges, only three independent variables could be used in the analyses to attain valid results. Using more than three predictors here could have distorted the regression analysis.

ANOVA^a

Sig.	F	Mean Square	df	Sum of Squares	Model 1
.027	3.26	3.01	3	9.020	Regression
		.923	65	59.965	Residual
			68	68.986	Total

a) Dependent Variable: Sentences reflect a lenient approach

Two comments need to be made here. First, the rather small proportion of variance might be due to the small variance in the dependent variable. As can be seen in Chart 1, there was a small variance in the judges' responses to this question: 83 percent of the responses were concentrated in three of the six response categories. This small variance may be a reason for the small variance explained by the model analyzed here. The relatively small number of respondents may also be a contributing factor.

Second, there were some other variables not measured in our survey that might have had an impact on the way judges evaluate the severity of prevalent punishment. Therefore, the three variables in the model here cannot have been expected to account for the large portion of the variance. Future research ought to explore for other variables and examine the effect on judges' assessment the level of punishment.

Given the abovementioned limitations, rather than focusing on the total percentage of variance accounted for by the model, we suggest focusing on the relative weights or relative importance of the variables in the model (β weights)⁷³ in accounting for the variance in responses. From this perspective, the tenure as a judge ($\beta = 0.40$) appeared to be the main predictor related to punishment assessment, suggesting that the longer the years as a judge, the more lenient the assessment of punishment is. The second predictor was court instance ($\beta = 0.29$) in which the second instance judges saw the punishment as more lenient than the first instance ones. The third factor, gender, was not found here to have any net effect at all on the assessments.

⁷³ β is a standardized regression coefficient (i.e., coefficient expressed in standard score) that reflects the net weight of each variable (accounting for that of the others) in the regression formula.

V. CONCLUSION

The main finding of this study is that the judges viewed the general level of sentencing as consistently lenient. This is surprising since one would expect that judges, who are the ones making the penalty decision, would view the courts' punishment as appropriate, not as considerably lenient. Moreover, our findings showed that the judges' perception of the leniency of criminal punishment was not just an isolated phenomenon. It was a consistent assessment. Over three consecutive surveys, each taken a decade apart, the Israeli judges repeatedly indicated that they regard the courts' sentencing as lenient.

These findings suggest that the courts system has no effective regulatory or internal control mechanisms to adjust or 'correct' the level of sentencing. Adjustments or corrections were not found even when extended over a long period of time, a finding that is quite surprising. To wit, the courts' system lacked (or failed to effectively maintain) an inherent control mechanism (homeostasis). The absence of such mechanisms calls for the introduction of regulatory homeostatic 'safe-guards,' either internal or external. Without such regulatory mechanisms, the level of punishment issued by the courts may diverge from the 'appropriate' sentencing level.

Furthermore, the results suggested that the judges themselves did not take action to adjust the level of sentences they believed to be too lenient. They were the agents who determined the punishment and its level and had the discretion to change it. They could have served as an 'adjusting mechanism' by issuing sentences that were more in line with what they themselves consider 'appropriate' punishment. i.e., issue more severe sentences when they perceive the general level of punishment to be lenient. Similarly, they could have issued sentences that were more lenient when they perceived the level of punishment to be overly severe. However, the findings here indicated that this was not done. It appears as if the judges consistently acted contrary to their own judgment: lenient sentences were issued while simultaneously assessing the general level of punishment to be considerably lenient.

Why would judges be reluctant to adjust their sentencing level to what they think is right? One possible answer to this question may be found in Posner's question: 'What do judges maximize? (The same thing as everybody else

does)⁷⁴ i.e. their own interests. Indeed, in a previous study we conducted that considerations of personal reputation might affect judicial behaviour.⁷⁵ Given this, we propose that although judges have the formal and structural independence to issue the 'appropriate' punishments, a judge's personal consideration of his reputation may underlie his reluctance to take action toward correcting lenient punishment.

In Israel, as in many other Western democracies, judges are given discretion and independence to make judgments as they see fit (within the law), free of external pressures. They are legally and structurally insulated from extraneous considerations and influences and are bounded only by the letter of the law. It is the judges' duty to impose appropriate punishment. What could drive judges to impose lenient sentences? Judges, as does 'everybody else,' may have their own interests in mind. For example, they may have an aversion to being overruled and reversed. Such reversals may damage the reputation as a competent judge.⁷⁶ Lenient (but not too lenient) sentences appear to be the best strategy for judges to minimize reversals. In general, the probability of the defendants appealing a decision is substantially higher than that of an appeal by the prosecution (the State). A judge can quite easily decrease the probability of reversal by imposing lenient sentences (but not too lenient). The probability of appeal under such a leniency policy will be diminished because the defendant may be reluctant to appeal fearing a harsher sentence, and the prosecution (given the punishment is not too lenient) may be reluctant to allocate the time and resources involved in an appeal, once a guilty verdict has been attained. Consequently, the probability of reversals is minimized, and with it, the probability of damaging the judge's reputation.⁷⁷ Hence, an optimal strategy for a judge to avoid reversal is to

⁷⁴ Richard A. Posner, 'What Do Judges and Justices Maximize? (The Same Thing Everybody Else Does)' (1993) 3 *Supreme Court Economic Review* 1-41.

⁷⁵ Moshe Bar Niv (Burnovski), Ran Lachman, 'Self-Interest in Judges' Time Allocation for Writing Judgments' SSRN: <http://ssrn.com/abstract=1641376>, or <http://dx.doi.org/10.2139/ssrn.1641376> accessed 16 July 2010.

⁷⁶ See, Christopher R. Drahozal, 'Judicial Incentives and the Appeals Process' (1998) 51 *Southern Methodist University Law Rev.* 469-503.

⁷⁷ Various studies claim that judges are influenced by the fear of reversal, see for example, Stephen J. Choi, Mitu Gulati, and Eric A. Posner, 'What Do Federal District Judges Want? An Analysis of Publications, Citations, and Reversals' (2012)

impose lenient penalties. Since this was not the focus of our studies, we have no current data to test this proposition. Further research is required in order to find an answer to this puzzling issue.

The leniency of punishment, as reflected by the judges' assessments over decades, ought to be a source of concern for the system. Such a continuously lenient level of punishment is likely to have a considerable impact on the behaviour of the various stakeholders in the justice system. Potential offenders, as well as potential victims of crime, may react to it by developing socially undesirable patterns of behaviour. If left unobserved and uncorrected, such a trend may result with grave consequences such as erosion of the public trust in the judicial system. Such erosion might undermine the whole system of judgement's foundations and legitimization.

28 *Journal of Law, Economics & Organization* 518-549; David E Klein, Robert J. Hume, 'Fear of reversal as an explanation of lower court compliance' (2003) 37 *Law & Society Review* 579-581.

APPENDIX

1. The Research Method

This study was a longitudinal study that examined, over a period of almost three decades, the perceptions of Israeli courts judges regarding the severity of punishment issued by Israeli courts. The study was based on the analyses of data collected in three consecutive surveys, conducted a decade apart in 1989, 1999 and 2010, among the judges presiding in first and second instance courts in Israel. In each survey, self-administered mail-questionnaires were sent to all the presiding judges in the first and second instance courts (Supreme Court judges were not included).

The three surveys included questions on a variety of issues and opinions related to work of the judges and the functioning of the judicial system. These were outside the scope of the present study. In all three surveys, however, the very same question was asked regarding their assessment of the level of punishment. The identical question allowed comparison of the judges' responses across the time period of nearly three decades.

The self-administered questionnaires were anonymous and the confidentiality of respondents' answers was promised. Judges were asked to fill the questionnaire and mail it back to the researchers in a pre-addressed and stamped envelope.

2. Survey Procedure

Prior to conducting the surveys, exploratory studies were conducted to investigate the phenomena we intended to study and find out the judges' views on these issues. These studies included a literature review as well as a number of in-depth interviews with judges (mostly retired judges or judges who had voluntarily left the bench before retiring) to get their perspective on the issues at hand. Based on these interviews, a first draft of the survey questionnaire was constructed. The draft was pre-tested by several (ex-) judges and their comments on it were integrated into a final draft of the questionnaire. To allow for over time comparisons, the questionnaires of the three surveys were largely identical except for one part in them relating to issues that were topical at the time of the respective surveys.

To each of the judges who were included in the survey, a personally addressed envelope was sent containing the questionnaire, the approval by the Chief Justice of the Supreme Court to distribute the questionnaire, and a stamped envelope with a return address. Since the private addresses of judges in Israel are confidential, questionnaires were sent to the judges by mail to their courts. Since the questionnaires were anonymous and the envelopes were not marked and had no signs identifying the sender, it was impossible to know who sent back his or her questionnaire or what a given judge answered. Hence each time, two weeks after the survey was sent out, a reminder was sent to all judges, emphasizing that those who had responded should not respond again. 49, 65 and 86 judges (respective to the surveys) sent back completed and usable questionnaires (see below).

3. The Respondents

The study populations were all the judges in Israel who presided in first and second instance courts at the time of the respective survey (not including the Supreme Court, Labour Court, Juvenile Court, etc.). When the first and the second surveys were conducted (1989 and 1999) 320 judges in total served in the first and second instance courts in Israel. The total number of judges at the time the third survey was conducted (2010) had increased to 528 judges.⁷⁸ Our aim was to survey the entire population of judges, not just a sample of them. Thus, each of the three survey questionnaires was sent to all judges in Israel. However, as would be expected, not everyone responded to the questionnaire. Consequently, the resultant sample was obtained of those who were kind enough to respond and send the questionnaire back to us: 49 in 1989, 65 in 1999 and 86 in 2010.

As these respondents do not constitute probability samples, each was examined to see if they represented the judges' population at the time. The characteristics of the respondents in each sample were compared to the overall characteristics of the judges' population. In all three samples, no statistically significant differences were found. In other words, the samples appeared to represent the population of judges in Israel well. Further, we tested whether the samples obtained were compatible with each other. We

⁷⁸ The number of judicial positions has increased in response to a shortage of judges in Israel.

compared a few characteristics of the responding judges in the three samples. For example, the average tenure in the legal profession and years in office (as a judge), were found to be similar (no significant differences) in the three samples. In the 1989 sample, the average tenure of the judges in the profession was 22.5 years, while it was 26.4 years in the 2010 sample. The average number of years in office as a judge was 9.9 years in the 1989 sample, compared to an average of 9.6 years in the 1999 survey and 11.6 years in the 2010 survey. To wit, over time, the three samples of judges were not different from each other. It appears, therefore, that the three samples (a decade apart) were compatible with each other as well as representative of the judge population in Israel.

Although representative of their respective populations (which were small to begin with) the absolute numbers of respondents in the 1989, 1999, and 2010 samples were rather small (49, 65 and 86, respectively). Hence, it cannot be concluded with high confidence from the findings that such conclusions unequivocally apply to all judges. Nevertheless, since surveys in which the judges themselves answer the questionnaires concerning their work are extremely scarce, we found it important to present the findings and treat them as indicating general trends and suggestive of possible implications.

4. Possible Response Biases

The small number of respondents relative to the population size could raise concerns that a self-selection bias may exist: i.e., respondents decide whether or not to respond to the questionnaire based on their interest (or disinterest) in the study topic. Consequently, the sample might not be random but biased by the judges' views on the research topic.⁷⁹ However, such a bias is not very likely in these surveys: the questionnaires included a large number of topics about the legal profession (such as the very large increase in the number of lawyers in Israel between 199 and 2010, and its implications for the profession), judicial (judging processes, decision making, etc.), the functioning of the judicial system, and more. It is unlikely that judges would pick out one question (the evaluation of the severity of prevalent courts' punishment) to decide based on it whether to continue filling it out or not. Secondly, as indicated above, other characteristics of the judges in the

⁷⁹ Earl Babbie, *The Practice of Social Research* (Cengage Learning 2011).

samples (e.g., distribution to districts, gender, etc.) were similar to their distributions in the judges' populations as a whole in Israel. Thirdly, we have interviewed several judges about their willingness to respond to the questionnaire, and these interviews showed that lack of time and work-overload were the main reasons for not answering the detailed questionnaire. Given all this, we suggest that there was no significant self-selection bias on the samples.

5. Measuring the Dependent Variable

The dependent variable in our study was the judges' perception of the severity of the courts' punishment. This was measured by a question asking the judges (in all three surveys) to respond to the following statement: 'Given your familiarity with the Israeli court system, would you say the punishments issued in Israel generally reflect a lenient approach?' The response range was: 1. Not at all; 2. To a small extent only; 3. Partially so; 4. To a considerable extent; 5. To a large extent; 6. To a very large extent. A few points should be made regarding the question and the choice of response range.

In order to assess the judges' evaluation of punishment, a single, straightforward question was used. Given the application of a self-administered survey method, the questionnaire had to be as succinct as possible in order not to put-off the respondents. This is common practice in public opinion polls and attitude-surveys where no face-to-face interview is held.⁸⁰ This practice is even more pronounced when respondents are highly professional and very sensitive to time pressures, such as was our case with the judges. In the pre-tests we conducted, the question regarding leniency was tested, and the judges participating in the pre-test indicated that the question was quite clear.

A second point is the definition of 'lenient'. Since perceived 'leniency' is an individual evaluation ('in the eyes of the beholder', so to speak), it had to be left to the individual judge to decide what he/she considered to be lenient and to what extent. Defining it for the respondent would entail 'imposing' the researchers' definition on the respondents and, therefore, is seldom

⁸⁰ E.g., Frederick J Gravetter, Lori-Ann B. Forzano, *Research Methods for the Behavioral Sciences*, (Wadsworth Publishing 3rd edn, 2009).

recommended in cases like ours. Furthermore, by law, judges are given the discretion to determine what the adequate punishment is for a given case (except for the few offenses where a mandatory punishment is determined in the law). Judges are trusted to use their best judgment. Hence, it is up to them to determine what level of punishment is 'inadequate' and to what extent. Also, they are human, so they act according to their own perceptions and, therefore, it is important to assess these perceptions. Given these considerations, the question presented asked them to assess whether the extent the punishments issued in Israel generally reflect a lenient approach not according to a pre-determined 'objective yardstick' but given their familiarity with the Israeli court system.

The response range chosen appears to overemphasize the leniency of punishment, because it is not 'symmetrical' in the response options (i.e., not symmetrical between 'lenient to a very large extent' on the one end and 'severe to a very large extent' on the other). The response range in our study was not symmetrical as it offered several response options regarding leniency and only a single answer of 'not at all lenient', and had no reference to the possibility that the punishments were harsh. We chose to regard the anticipated distribution of answers to this question to be skewed to the side of leniency. In the literature review preceding the study, we had found that all the previous polls and studies found that only a small percentage of the public evaluated punishment as harsh; the vast majority of respondents indicated that the punishments were lenient.

Furthermore, in preliminary interviews we conducted with judges and lawyers, the interviewees overwhelmingly stated that the punishment was too lenient, with not even one suggesting that it was too severe. In addition, the preliminary results of the pre-test study showed that the respondents did not select the response of 'severe punishment'. Hence, this response alternative was omitted. Designing the frequently used 'symmetric' response range for this question would have distorted the skewed nature of the response distribution.⁸¹

Our preference was to create a scale more sensitive to tapping differences on the side of the more frequent response in order to better distinguish between

⁸¹ *ibid.*

various dimensions of leniency. A symmetric range or scale of, for example, five possible answers, ranging from: 1. very lenient to 5, very harsh, would have meant allocating two response options on the severe side of the range, two on the lenient side, and one middle or neutral option. In cases such as the present one, most responses would have been given on the too 'lenient' side of the scale with few if any responses on the two options allocated to the severe side of the scale. Two response options are not a sensitive enough scale to differentiate perceptions within the 'lenient' side. At the same time, the two 'harsh' punishment responses would have been left with almost no response. A concrete example of this situation can be seen in the findings of a study conducted in England, shown in Table 1 above. There, interviewees were presented a symmetrical response-scale ranging between strict and lenient punishment, and the results showed that in three consecutive surveys only three percent (combined) of the respondents responded to the two options of 'strict' and 'very strict' punishment. In contrast, three-quarters of responses were concentrated on the two 'lenient' options. Given such results, it is: a. not possible to get a fuller range of answers regarding the real perceptions; and b. hampers the ability to analyze and better understand the assessment of punishment by the public. Indeed, the actual distribution of responses found and reported here (for example, Chart 1) is concentrated on the 'lenient' side of the scale while there were very few that answers of 'not at all lenient.' Thus, it supports our preference of the asymmetric response range over the symmetric one.