How German Labor Courts Decide: An Econometric Case Study

Helge Berger
Free University Berlin and CESifo

Michael Neugart
Free University of Bozen/Bolzano

Abstract. *Courts are an important element in the institutional framework of labor markets, often determining the degree of employment protection. German labor courts provide a vivid example in this regard. However, we know relatively little about court behavior. A unique dataset on German labor court verdicts reveals that social and other criteria like employee characteristics, the type of job, local labor market conditions and court composition influence court decisions. At least as striking is that workers' chances to win depend on where and when their cases are filed. This generates considerable* ex ante *uncertainty about outcomes.*

1. INTRODUCTION

Labor courts are an important element in the institutional framework of labor markets. Most employment protection indicators focus on laws, but in many countries it is labor courts that determine the degree of employment protection. However, in contrast to legal norms which are relatively easily documented and verified (OECD, 2004), little is known about the behavior of labor courts.

German labor courts provide a vivid example of the importance of the judicial branch. Employment protection in Germany is rooted in a number of laws ranging from the Civil Code (*Bürgerliches Gesetzbuch*) to the Protection against Dismissal Act (*Kündigungs schutzgesetz*), and the Works Constitution Act (*Betriebsverfassungsgesetz*). However, these laws set out general principles rather than specific rules and in practice German labor courts determine the degree of employment protection by implementing, interpreting and often developing the legal framework (Berger, 1998; Goerke and Pannenberg, 2010; Grund, 2006). In a three-level judicial system, the task of developing the rule governing employment protection falls mostly on the second (*Landesarbeitsgerichte*) and third (*Bundesarbeitsgericht*), while implementation takes place at the first (or lower) level (*Arbeitsgerichte*), which handles the brunt of labor court cases.

Contract disputes over dismissals are behind about half of all cases handled by lower-level labor courts, and estimates suggest that between 11% and 27% of all dismissed workers go to court (see Höland *et al.*, 2007; Jahn and Schnabel, 2003; Pfarr *et al.*, 2005). Here judges rule, on a case-by-case basis, whether dismissals are justified according to a number of law-anchored but mostly court-developed criteria. With regard to dismissals based on economic reasons, labor courts regularly check against 'social criteria.' Among these criteria is the age of a worker disputing

his or her dismissal, his or her tenure, and obligations to support dependents.[1] If found lacking along these lines, dismissals can be revoked or courts will determine that employees should receive a severance payment from the firm.

The present paper provides an empirical case study of German labor court decisions and their determinants. It is based on a unique dataset illustrating the similarities and differences of court behavior across Germany, akin to a natural experiment in which a researcher would have assigned the same case in random fashion to different courts. The data summarize the decisions and key characteristics of 221 individual cases handled by 33 different lower-level labor courts in 12 German states between August 2003 and September 2006. The underlying case characteristics invite analysis of cross-court heterogeneity as all cases were filed by employees of the same national electronics and media retail chain, the causes of dismissal were virtually identical across the sample, the retailer was represented by the same law firm and lead attorney throughout, all outlets were covered by either one or two work councils collaborating closely, and all cases occurred within a rather short time interval.

The results include a number of surprising facts. While social criteria anchored in the legal framework indeed play a role in explaining court decisions, other employee characteristics (such as gender or the presence of a union attorney), the nature of the job held by employees (including the salary level), and local labor market conditions (unemployment) also matter. While some of these results can be reconciled with the idea of social criteria in a broader sense, others cannot. This seems to point to a more discretionary side of labor court decision-making, with outcomes that may not necessarily be intended by lawmakers. This certainly also holds for indications that the probability of winning a case seems to be influenced by the gender composition of labor courts. At least as striking is the very strong impact of fixed time and court effects, which suggest that workers' chances to win depend systematically on where and when their cases are filed. Finally, while these findings indicate a certain predictable regularity in court behavior, considerable *ex ante* uncertainty about outcomes remains.

Some of these findings reflect results reported in earlier studies of court behavior. For instance, Franz (1994) discusses anecdotal evidence suggesting that German labor courts have trended toward more employee-friendly decisions after the surge in unemployment rates in the late 1970s. Bertola *et al.* (1999) review international evidence pointing in a similar direction. And, in a study close in spirit to our approach, Ichino *et al.* (2003) illustrate heterogeneity in labor court behavior across Italy and establish a link between labor market conditions and court decisions. An advantage of their analysis compared with ours is that their data allow some modeling of the decision of workers to file a court claim. A disadvantage is that the causes for dismissal are not naturally restricted and case dates are spread over a wider time period, which limits comparability and increases the risk of omitted variables. Our study robustly confirms that courts are more likely to rule in favor of employees in regions with depressed labor markets even after controlling for a host of individual characteristics of the claimants, defendants and judges involved in a particular case.

---

1. On 'social criteria,' see, inter alia, Däubler *et al.* (2009), Hromadka and Maschmann (2002), or Kittner and Zwanziger (2001).

Other related literature includes research stressing gender issues in court behavior because female judges might be more attuned to problems faced by female employees – see, for instance Boyd *et al.* (2009) on US courts. There is also a set of papers exploring the severance payments granted by German labor courts in light of social criteria, tax law, and other possible determinants (see, e.g., Goerke and Pannenberg, 2009 or Jahn, 2005, 2009, respectively). In addition, it has been argued that court decisions could be subject to political influences. For instance, there is some evidence that US and German judges react to societal preference swings or share certain biases with the politicians appointing them (Berger and Neugart, 2008; Hanssen, 2004). Finally, Hefeker and Neugart (2010) stress the link between the behavior of labor courts and labor market policies and show that the degree of court discretion (and the resulting uncertainty) influences regulatory activity in a panel of OECD countries.

There is reason to believe that labor court activity and, by extension, the determinants of labor court activity, influence labor market performance. At a theoretical level, Stähler (2008) argues that it may particularly be uncertainty about labor court decisions which impacts employment. Developing the Pissarides (2000) matching model to allow for shirking along the lines of Shapiro and Stiglitz (1984), he shows that judicial mistakes can increase unemployment.[2] For example, a firm might claim that the worker was dismissed for behavioral reasons although the dismissal was actually operational and the court, for whatever reason, wrongly denies severance payments. Judicial mistakes introduce additional income uncertainty into workers' income, which increases the wages a firm has to pay in order to avoid shirking and, ultimately, lowers employment.[3] Empirically, there is some evidence that labor court decisions indeed influence labor markets outcomes. Autor (2003) and Autor *et al.* (2004, 2006) present results for the United States, and Berger and Neugart (2008) report a significant positive relation between labor court activity and unemployment in Germany after controlling for the endogeneity of court activity.[4]

Any analysis of court behavior needs to be concerned about decisions by workers and firms taken before the filing of a case because they can, at least in principle, influence sample selection and empirical results. In our case study, this could include the conduct of workers before the dismissal, the firm's decision to fire workers, and the decision by workers to take the case to court. However, while we know that there were about 2,000 dismissed workers (with and without re-employment offers) overall, we only have information about the 221 workers that went to court, which prevents us from exploring this route. That said there are reasons to believe that selection bias plays no significant role in our dataset. Empirically, there seem to be no discernible differences in sample characteristics between our data and comparable studies (see Section 2). Additionally, the firm's decision was dictated by an overarching downsizing rationale (through the outright closing of outlets or downgrading to simpler self-service operations) and not driven

---

2. See Galdón-Sánchez and Güell (2003) for an earlier exploration of judicial mistakes in a shirking framework. Goerke (2002) discusses the advantages of severance pay over pure firing costs from an allocative perspective.
3. Huang *et al.* (2009) argue that, at least in part, the mistake made by courts could also have a deterministic component that could be correlated, for instance, with unemployment. This would open another feedback loop between court behavior and the real economy.
4. See also earlier findings by Berger (1998) and Berger and Danninger (2006).

by individual worker-specific considerations, which speaks against a systematic selection bias on the firm's side. And even if workers had perfectly anticipated the event, there would have been little or no reason to systematically adjust their behavior. As to the question whether workers with certain attributes were more likely to file their case than others, the answer depends. With respect to the personal characteristics, self-selection would only be an issue to the extent that the composition of the workforce differed systematically across regions, which does not seem to be the case. This leaves the regional labor market conditions that indeed vary across the labor court locations. However, while this could potentially upward-bias the estimated impact of, say, regional unemployment rates on court outcomes, it should not distort the qualitative results. Moreover, available evidence from comparable studies suggests that any such effects might be very small. Goerke and Pannenberg (2010), using a Heckman (1976) procedure to control for the possible self-selection of workers going to court, report little or no evidence of an endogeneity bias.

The remainder of the paper is organized as follows. Section 2 describes the data in greater detail. Section 3 reports the results of a more extensive econometric analysis, whereas Section 4 provides a discussion of the robustness and representativeness of results. Finally, Section 5 concludes.

2. DATA

Our dataset comprises 221 written decisions by lower-level labor courts on dismissals. About 3/4 of these claims have been filed by employees aiming to nullify an operational dismissal (*Beendigungskündigungen*) by the employer, and the remaining 1/4 of cases concerned employees seeking to cancel an operational dismissal that was combined with a conditional re-employment offer under a new contract (*Änderungskündigungen*), usually implying a significant downward change in job title, working conditions and pay. While the underlying legal concepts differ, both types of dismissals turn out to be more or less equivalent in empirical terms. This is hardly surprising. Almost by definition, the employees taking their cases to the courts were also the ones that found the conditions attached to the offered re-employment unattractive. That is, from their perspective, the dismissal combined with conditional re-employment was equivalent to a straightforward dismissal. In what follows, we will present our findings based on the full sample of court decisions.

In a typical case, a worker received a dismissal note or had to accept a new job assignment. This worker brought the case to its labor local court asking to nullify the dismissal for social or other reasons. While the court process routinely incorporates a conciliation phase (*Güteverhandlung*), our sample does not include settlements, reflecting the firm's decision to seek a court decision in every case. Accordingly, a 'win' by the worker typically implied that the court annulled the dismissal or job assignment. In the empirical application, we coded a court decision as 'won by employee' if courts sided with workers, and as 'not won by employee' otherwise. The latter category encompasses all dismissals fully approved by the court, including, for example, rulings that no severance payments were imposed on the firms.

As Table 1 reveals, about 62% of all cases in the sample were won by the workers, but there is considerable regional variation. For example, at labor courts located in

Table 1 Distribution of cases by labor court and winning employees

Location of labor court	Total number of cases	Cases won by workers	%
Bamberg	2	2	100.0
Bayreuth	2	2	100.0
Berlin	67	36	53.7
Brandenburg a.d.H.	3	1	33.3
Braunschweig	2	0	0.0
Bremen	3	2	66.7
Chemnitz	1	0	0.0
Darmstadt	1	1	100.0
Deggendorf	1	1	100.0
Dessau	4	0	0.0
Düsseldorf	1	1	100.0
Freiburg	16	15	93.8
Halberstadt	2	1	50.0
Halle	11	6	54.5
Heilbronn	8	1	12.5
Herne	1	1	100.0
Karlsruhe	6	6	100.0
Landshut	1	1	100.0
Leipzig	1	1	100.0
Ludwigshafen	18	2	11.1
Magdeburg	2	2	100.0
Mainz	6	3	50.0
Mannheim	29	29	100.0
München	6	2	33.3
Naumburg	1	1	100.0
Neubrandenburg	1	0	0.0
Nürnberg	5	3	60.0
Regensburg	6	4	66.7
Solingen	1	0	0.0
Stuttgart	2	1	50.0
V.-Schwenningen	5	5	100.0
Weiden	5	5	100.0
Zwickau	1	1	100.0
Total	221	136	61.5

Freiburg and Mannheim, employees won 100% and 94% of the decisions, respectively, while that ratio was about 54% in Berlin and Halle, and the courts in Dessau and Braunschweig decided against employees in every case they handled. The regional distribution involves a total of 33 courts in 12 (out of 16 overall) German states. The average number of decisions is about seven, but the median is two, and 11 courts, including, for instance, Chemnitz, Darmstadt and Deggendorf, saw only one case. At the same time, the number of labor court cases decided in Ludwigshafen, Mannheim or Berlin ranges from 18 to 67.

Turning to the time domain, all court decisions were made within a fairly short period. The first labor court decision in the sample dates from 20 August 2003 and the last from 7 September 2006. Generally, the case loads at local labor courts were increasing at a steady rate during this time period, and all cases were dealt with on an individual basis.

Table 2 Summary statistics of explanatory variables

Variable	Observations	Mean	Median	Minimum	Maximum
Tenure (years)	216	8.62	7	1	39
Age (years)	176	40.5	40	24	62
Children[a] (dummy)	221	0.29	.	.	.
Married[a] (dummy)	221	0.41	.	.	.
Gender (dummy)	221	0.37	.	.	.
Citizenship[a] (dummy)	221	0.95	.	.	.
Union attorney (dummy)	221	0.80	.	.	.
Gross monthly wage (euro)	207	2,220	1,950	825	12,867
Majority female (dummy)	216	0.17	.	.	.
Same gender (dummy)	216	0.65	.	.	.
Regional unemployment rate (%)	221	13.7	11.2	6.6	22.0

[a] Proxy variables, see main text.

In addition to the outcome, the written court decisions include a wealth of further information (see Table 2). These case characteristics are distributed fairly evenly across courts – for example, they do not differ significantly between Berlin, Freiburg, Ludwigshafen and Mannheim and the remainder of the sample

In terms of relevant socioeconomic employee characteristics, the court documents list the job *tenure* of the employee. Average tenure is nine years, with the shortest reported tenure being one and the longest 39 years. Based on the information we have for *age*, the average worker in the sample is about 40 years old; the youngest is 24, while the oldest is 62. As to *gender*, about 37% of workers are female. All workers had legal representation, but 80% were represented by *union attorneys*. The latter two variables are coded as dummy variables, taking on the value of one if employees are female or have union-supported legal representation, respectively.

The socioeconomic information discussed so far is available in most cases and quite straightforward to use, but the court documents are slightly less complete in other regards. For instance, the nationality of workers is not explicitly identified, forcing us to construct a very crude proxy variable for German *citizenship* (computed as a dummy variable) based on the family names noted in the court decisions. Based on this very rough proxy measure, about 95% of workers in the sample could have been German citizens. There is also a lack of information regarding the marriage or family status of workers. Here, too, we constructed proxy indicators based on the assumption that the court documents would have included this information if it had ultimately been relevant for the court decision. Consequently, we generate dummy variables taking the value of one if the court document explicitly reports that the worker has one or more children or is married, respectively. Based on this procedure, about 41% of all employees are classified as *married* and about 28% as having *children*.[5]

The data also include information on the jobs held by the workers going to court and their salaries. The *gross monthly wages* cover a range from 825 euro to (an

5. This procedure increases the number of observations for the variables *children* and *married* which are 109 and 143, respectively, to 211. It also reduces the share of non-zero observations from about 63% to 41% for the variable *married* and from about 58% to 29% for the variable *children*. We will discuss the robustness of results with regard to these changes below.

exceptional) 12,867 euro, with a mean of 2,250 euro and a median of 1,950 euros. To capture the type of *job* held by workers before the dismissal under dispute, we construct a set of dummy variables indicating the five professional categories occurring in the sample. These categories include: (1) sales and technical support, (2) cashiers, assistants and accountants, (3) storage workers, (4) managerial positions in sales or accounting, (5) other managerial positions (see Appendix A for details).

In addition, the dataset also contains information on the gender composition of labor courts. Each court comprises three judges, one principal judge and two secondary lay judges. The dummy variable *majority female* indicates when at least two of the three judges of a particular labor court were female, which is the case in about 17% of the cases in the sample. Moreover, to capture any interaction between the gender composition of courts and plaintiffs, the dummy variable *same gender* indicates whenever worker and the majority of judges were either female or male. This holds in about 65% of all observations.

We are also interested in a measure of the regional labor market performance. Since cases get assigned to specific labor courts based on the location of the workplace, we can match court location and county-level (*Kreis*) regional unemployment rates taken from the records of the German Employment Agency (*Bundesagentur für Arbeit*). The variable *regional unemployment rate* averages 13.7% over the regions and time periods covered, with a low of 6.6% and a maximum of 22%.

It is interesting to ask how broad a picture our case study can draw. There is some evidence that suggests that the sample characteristics are fairly close to other studies in the field. For example, Goerke and Pannenberg (2010), who look at occurrences of severance pay in a sample of almost 3,000 lay-offs in a large and widely used German panel dataset (*Soziooekonomisches Panel*) between 1991 and 2006, report broadly similar sample averages in terms of age, citizenship and regional economic conditions.[6] Even if sample characteristics were to differ, our results should remain relevant because of the random nature in which cases were allocated to courts (the natural experiment alluded to in Section 1).

3. ECONOMETRIC ANALYSIS

In what follows, we take a more systematic look at the determinants of the labor court decisions contained in our case study. Table 3 presents the results from a series of probit models explaining the likelihood of workers to win their cases in front of labor courts. The dependent variable is one if the worker has won the case and zero if the worker lost.

Column (1) reports the results of a simple model including only court- and (quarterly) time-fixed effects, both yielding jointly significant coefficients. This suggests a fair amount of court-specific variation in workers' probability of winning or regional bias, mirroring the observations in Table 1.[7] The strong presence of

---

6. Looking only at West German cases, the average age of workers in our sample (in the Goerke and Pannenberg, 2010 sample) is 40.4 (40.4%), 92 (86%) are German citizens, and the average regional unemployment rate is 9.9 (10.0%) after adjusting for differences in regional coverage.

7. Note that including court-fixed effects reduces the number of usable observations to just below 150 because for a number of labor courts with only few cases the fixed effects completely explain the outcome and these observations were subsequently dropped from the sample.

Table 3 Probit regressions for employee winning

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Tenure		0.04					−0.06	−0.02
		(0.273)					(0.088)	
Age		−0.03					−0.02	−0.01
		(0.000)					(0.128)	
Children		0.20					0.52	0.20
		(0.243)					(0.030)	
Married		0.18					−0.07	−0.03
		(0.412)					(0.760)	
Gender			0.44		0.53		2.39	0.76
			(0.043)		(0.017)		(0.000)	
Citizenship			−1.44				−5.65	−0.70
			(0.020)				(0.000)	
Union attorney			0.99				2.05	0.55
			(0.055)				(0.002)	
Gross monthly salary				-1.4×10^{-4}			-3.4×10^{-4}	-10^{-4}
				(0.206)			(0.018)	
Job dummies				Yes			Yes	.
				(0.001)			(0.000)	
Majority of female judges					0.21		0.22	0.09
					(0.331)		(0.467)	
Same gender					0.34		1.65	0.58
					(0.013)		(0.000)	
Regional unemployment						1.06	3.91	1.56
						(0.011)	(0.006)	
Court location dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	.
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	
Quarterly time dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	.
	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	
Number of observation	149	114	149	134	148	149	104	
Pseudo log likelihood	−80.28	−59.72	−76.23	−65.92	−78.53	−77.18	−35.39	
Pseudo-R^2	0.22	0.24	0.26	0.29	0.23	0.25	0.51	
Fit overall	109 (149)	82 (114)	112 (149)	105(134)	111 (148)	108 (149)	92 (104)	
Fit worker won	55 (75)	46 (59)	56 (75)	55(68)	60 (75)	58 (75)	49 (55)	
Fit firm won	54 (74)	36 (55)	56 (74)	50(66)	51 (73)	50 (74)	43 (49)	

Notes: Numbers in parenthesis in the upper half of the table report the *p*-values for individual coefficients or for Wald tests of joined significance. See main text for a discussion of the goodness-of-fit measures. All regressions use the STATA clustering option based on court location.

common time variation could indicate changes in the common legal framework, but potentially also the economic or political setting within which courts acted.

The pseudo-R^2, a broad indicator of statistical adequacy for probit models, suggests that the model explains a relevant share of the variation. Another way of illustrating the degree to which the model can explain the labor court decisions is to calculate the percent of correctly predicted cases (Wooldridge, 2006). Suppose employers or workers were to use the model presented in column (1) to predict the outcomes of trials, how many cases would be predicted correctly? For that purpose, we define a binary variable that takes the value of one if the predicted probability is larger or equal a threshold and zero otherwise. If actual and predicted outcomes are the same, the prediction is correct. For the cases where the actual outcome is one and the predicted outcome is zero and vice versa, the prediction is wrong. Applying a threshold of 0.5 (the robustness of the threshold is discussed in Section 4), the last three rows report the number of correctly predicted outcomes in relation to the overall cases given in parentheses, and for the subsets of worker or firms winning. We find that more than 2/3 of the outcomes are correctly predicted by model (1).

The regression reported in column (2) asks to which extent social criteria influence court decisions in addition to court-fixed and time effects. Here, as well as in all following regressions, we compute standard errors and *p*-values allowing for possible clustering of errors at the local labor court level. Despite the fact that *tenure* is seen as relevant among social criteria in the legal literature, our results indicate only a weak association with court decisions. While an increase in tenure seems to be linked to a higher probability of winning, the effect is not significant at conventional levels. In contrast, *age* shows a significant negative impact on the probability of employees winning, suggesting that – on balance – the courts in our sample tended to view older workers less eligible to social protection than younger workers. There is no evidence that the presence of *children* significantly increases the probability of a worker winning his or her case against a dismissal, and also the estimated positive coefficient for *married* remains insignificant.

There is, however, evidence that the social criteria variables are significant as a group.[8] Employee characteristics tend to be correlated. For instance, older workers will, in general, have higher tenure and be more likely to be married and have children. Their joint significance supports the view that social criteria do play a role for labor court decisions, and suggests their impact is best viewed group-wise rather than individually, perhaps reflecting some of the ambiguities about their relative weight discussed in the legal literature (see, e.g., Däubler *et al.*, 2009; Hromadka and Maschmann, 2002; or Kittner and Zwanziger, 2001).

Other employee-related case characteristics might also influence labor court decisions, even though the legal basis for this is a little less clear than in the case of social criteria. The underlying legal framework does not, for instance, suggest that the *gender* of the employee or its *citizenship* should play a role – quite to the contrary, an argument could be made that court proceedings should be 'blind' to such factors. At the same time, we would expect the presence of a *union attorney* to matter simply because good representation will generally boost the chances of winning court proceedings in front of a labor court. Column (3) shows our results, following the now familiar pattern of adding variables to the full set of court- and time-fixed effects. Somewhat surprisingly, the other employee-related case characteristics seem to matter

---

8. The hypothesis of *tenure*, *age*, *children* and *married* being jointly zero is rejected with a *p*-value <0.01.

to a significant degree. For example, we find that *citizenship* has a significant negative impact, even though it is important to keep in mind that the variable offers only a very crude proxy for German citizenship. Also *gender* has a significant positive impact, which seems to suggest female employees stood a somewhat better chance of winning their court cases than their male colleagues. The results for representation by a *union attorney*, while going in the expected direction, are somewhat weaker.[9]

Turning to job-specific variables, the legal background provides little or no help in formulating expectations on their impact on labor court decisions. We can speculate that a higher predismissal *gross monthly salary* or having held a *job* falling into a particular category might be associated with the ability to obtain better legal counseling. At the same time, a hypothetical court with a social agenda might also take its cues from these case characteristics. Column (4) shows that, while the estimated parameter on *gross monthly salaries* is insignificant, the *job* variables are jointly significant. Among the different dummy variables, 'other managerial-type positions' has the most significant individual impact, with the positive estimated coefficient suggesting that workers somewhat higher up the firm's hierarchy may have held an advantage when going to court.

Certain time-varying court characteristics might also influence outcomes. Column (5) of Table 3 shows that a *majority of female judges* has a positive but insignificant effect on the winning probability of employees. However, the interaction variable *same gender* – a variable that is one whenever both the worker and the majority of judges are either female or male – has a significant positive impact on the winning probability of workers. The result seems to suggest that labor courts operating under a majority of female judges are more likely to decide in favor of employees when the employee is female than when the employee is male. Similarly, a majority of male judges seems to have been more inclined to decide in favor of a male than of the female worker.[10] These findings are interesting but not easily explained. A simplistic approach would assume that judges are subject to an intrinsic bias, which would lead them to giving preferred treatment of their own gender. Other rationales discussed at some length in US studies of court behavior include that female judges might seize opportunities to undo (and, by extension, prevent) gender discrimination in the labor market or simply may bring to bear shared professional expertise and experience with female workers. However, it is difficult to differentiate between these or other explanations based on the available information. As Boyd *et al.* (2009) report, empirical results in the US literature are ambivalent, with some studies finding evidence of gender-specific behavior, others reporting mixed results and some no effects. See, for instance, Brudney *et al.* (1999), Davis *et al.* (1993), Giles *et al.* (2001), Peresie (2005), Segal (2000), or Sunstein *et al.* (2004).

In column (6), we explore the possibility that regional economic conditions might influence labor court decisions. The *regional unemployment* rate has a significant positive effect on the winning probability of employees. Given the presence of regional (i.e. court-) and time-fixed effects in the model, *unemployment rate* is likely to pick up any regional difference in the cyclical economic variation. This suggests that judges see a heightened need to protect workers when local economic conditions are deteriorating faster than elsewhere.

---

9. The hypothesis that all three variables are jointly zero is rejected with a *p*-value <0.01.
10. Because *same gender* is an interaction variable, the model also includes *gender* as additional control. All three variables are jointly significant with a *p*-value <0.01.

Column (7) incorporates the complete set of explanatory variables. Data restrictions reduce the available sample by about one-third compared with the partial models discussed so far, making a direct comparison of the results difficult. But the overall impression is that the findings are generally very robust. For instance, we can still not reject the joined significance of the social criteria variables, other employee characteristics remain relevant, and so do other job-related variables and those modeling gender issues and labor market influences.[11] If anything, the results are showing statistically stronger results. The model explains 92 out of 104 decisions overall.

How large are these effects in economic terms? To answer this question, the last column in Table 3 reports the marginal impact of each variable in model (7) evaluated at sample means. A number of the computed quantitative effects are quite small. Somewhat surprisingly, this holds particularly for some of the variables associated with social criteria and other employee characteristics. For instance, increasing *tenure* or *age* by one year from their mean values decreases the probability of a winning worker by as little as 1 and 2 percentage points, respectively. Similarly, being *married* has only a minimal effect on winning (a married worker's winning probability is about 3 percentage points higher than an unmarried worker's), and an increase of the *gross monthly salary* by 1,000 euro increases the likelihood of winning by as little as 0.1 percentage points.

Other variables have stronger effects. Among the social criteria variables, this holds true for the dummy variable *children*, where a childless worker has a 20 percentage points lower probability of winning than a worker where the court documents indicate that children are present. As to other characteristics, the presence of a *union attorney* improves the probability of winning by as much as 55 percentage points, and the effect of *citizenship* is even more pronounced, with workers that – according to the proxy variable used – could have a non-German background having a 70 percentage points larger likelihood of winning. The marginal effects of *gender*, *majority of female judges* and *same gender* are calculated based on a linear approximation adding up the marginal effects suggesting that, as a group, these case characteristics matter quantitatively, with the marginal effect of *gender* being in the order of 20 percentage points when the majority of judges is male, and the marginal effect of the *majority of female judges* variable being in the order of − 50 percentage points for male claimants.[12] Lastly, the marginal effect for the regional unemployment rate indicates that the slope of the probit function is larger than one at sample means. An increase in the unemployment rate by 1 percentage point would raise the likelihood of a worker winning by about 150 percentage points. The impact reading is strong but, as discussed earlier, should not be taken literally as a link from unemployment to court decisions – rather it could suggest a sensitivity of court decision to changes in the general economic well-being of a region.

Overall, a number of results are worthwhile stressing. Social criteria do matter, statistically and economically (if not all to the same extent), for court decision even

---

11. The joined *p*-value of the social criteria variables is ≤ 0.01, and so is the joined *p*-value for other employee characteristics.
12. See Ai and Norton (2003) and Norton *et al.* (2004), among others, for a general discussion of the difficulties involved in calculating these effects. We opted for a linear approximation as otherwise additional assumptions on the size of the other explanatory variables would have had to be made and furthermore the variance in the size of the effects across model specifications seems to outweigh any additional accuracy one could hope to gain from taking into account the non-linearity of the underlying probit model.

in a comprehensive model that includes a number of additional controls. This should be welcome on normative grounds in the sense that it provides confirmation that labor court decisions reflect the legal framework stressing these criteria.

There are several other case characteristics that labor courts seem to be taking into account in addition. This group of variables includes, among others, court- and time-fixed effects, regional labor market conditions, and the gender composition of the court. In general, these regularities are harder to square with the legal framework that one would expect labor courts to operate within. This could raise questions about the desirability of the apparent discretion labor courts enjoy in their decision making.

There also seems to be considerable uncertainty about court outcomes from an *ex ante* perspective. A number of considerations come into play. Looking at the pseudo-R^2 and goodness-of-fit measures, the comprehensive model in Table 3 clearly has larger explanatory power than the partial models discussed earlier. Ninety-two out of 104 cases or 88% of outcomes are predicted correctly. However, the goodness-of-fit does not look dramatically higher than in the simple models that include only time-fixed and court-fixed effects. There 109 out of 149 cases or 73% are predicted correctly. Moreover, a model including only time-fixed effects predicts 69 out of 104 cases or 66% if we base the sample on the comprehensive model. Arguably, even if workers had been able to anticipate perfectly the court-based fixed effects (that is, if they had been fully aware of any exogenous bias in their favor), they would have had a hard time anticipating the trend of winning probabilities across courts captured by the time-fixed effects. This also holds for some of the other explanatory variables, such as the courts' composition, which are not necessarily known *ex ante*.[13] Therefore, while the econometric results reveal a fair amount of predictable regularity in court behavior, considerable outcome uncertainty remains.

4. ROBUSTNESS OF RESULTS

Our findings are fairly robust along a number of important dimensions. First, to establish the sensitivity with regard to outliers, we re-estimate each model in Table 3 dropping ten observations in a moving time window starting from the first decision in our sample. For example, for the full model in column (7) of Table 3, which was estimated on 104 observations, the exercise produces 94 estimated coefficients for each of the explanatory variables (not reported). While the exercise can quickly exhaust degrees of freedom, the results generally compare well with the full sample estimates. Second, as already noted, our results are robust with regard to the approximations underlying the variables *children* and *married*. Third, the pooling of dismissal data for cases with and without an associated offer of re-employment does not seem to have a large impact on our results. While we cannot run the full model (7) for lack of observations, models (1)–(6) produce qualitatively similar results excluding dismissals with re-employment offers.[14] Fourth, changing the threshold values used to calculate the goodness-of-fit measures within the plausible interval between 0.4 and 0.6 does not alter outcomes by much.

---

13. Court composition can change. And where there are multiple lower-level courts at one location, cases are allocated randomly. For instance, in Berlin, 13 different principal judges handled the cases, and in Mannheim, Ludwigshafen and Freiburg, 5, 4 and 4, respectively.

14. There are some indications that *married* and *unemployment* play a somewhat larger role in models (2) and (6) when using the restricted sample.

5. CONCLUSION

Courts are an important element in the institutional framework of labor markets across the world, often determining the degree of employment protection through their interpretation and development of the law. German labor courts, which act in a legal framework putting some but not too many restrictions on their behavior, are an interesting example in this regard. Yet, surprisingly little is known about their behavior. A new dataset helps to fill some of this gap. The data include decisions and detailed information about 221 labor court cases handled by 33 different labor courts in 12 German states between 2003 and 2006. All cases were filed by employees of the same national electronics and media retail chain and, on the firm's side, were handled by the same law firm. The data document what is close to a natural experiment, illuminating differences and similarities of labor court behavior across Germany.

A number of noteworthy results emerge. Somewhat reassuringly, the labor court decisions captured in the sample reflect some of the more specific elements of the legal framework. Courts take into account so-called social criteria (such as age, tenure, or the presence of children) stipulated by law. At the same time, however, courts also show significant discretion in their decision-making by systematically reacting to other case characteristics. This group of variables includes, for instance, job types, regional unemployment conditions, and time- and court-fixed effects. The fixed effects suggest that workers' chances to win depend systematically on where and when their cases are filed. In addition, there are indications that court decisions are influenced by the courts' gender composition and the gender of the worker who filed the case. Some of these findings are surprising and could raise questions about the desirability of the discretion enjoyed by labor courts. Lastly, there is uncertainty about court outcomes from an *ex ante* perspective. While the econometric results show a fair amount of predictable regularity in court behavior, considerable ambiguity remains. The theoretical literature suggests that uncertainty of this type could have negative impact on employment.

The results should be taken with a pinch of salt because of the case-study nature of the empirical approach. To the extent that our findings can be generalized, they seem to imply that German labor market courts could influence labor market outcomes through the discretionary nature of some of their decision making and limitations to their predictability.

APPENDIX A

Table A1 Job categories

Category	Examples (job titles in German)
1. Sales, technicians	Fachberater(in), Verkäufer(in), Radio- und Fernsehtechniker(in)
2. Cashiers, assistants, accountants	Verkäufer(in) mit Kassentätigkeit, Kassierer(in), Sekretärin, Personalbearbeiter(in), Sachbearbeiter(in)
3. Storage workers	Lagerarbeiter(in)
4. Manager sales or accounting	Verkaufsleiter(in), Leiter(in) Controlling
5. Other managerial positions	Revisor, Disponent, Personalentwickler, Assistent der Betriebsleitung, Aktionsmanager

ACKNOWLEDGEMENTS

REFERENCES

Ai, C. and E. C. Norton (2003), 'Interaction Terms in Logit and Probit Models', *Economics Letters* 80, 123–129.

Autor, D. (2003), 'Outsourcing at Will: The Contribution of Unjust Dismissal Doctrine to the Growth of Employment Outsourcing', *Journal of Labor Economics* 21, 1–42.

Autor, D., J. J. Donohue and S. J. Schwab (2004), 'The Employment Consequences of Wrongful-Discharge Laws: Large, Small, or None at All?', *American Economic Review (Papers and Proceedings)* 94, 440–446.

Autor, D., J. J. Donohue and S. J. Schwab (2006), 'The Costs of Wrongful–Discharge Laws', *Review of Economics and Statistics* 88, 211–231.

Berger, H. (1998), 'Regulation in Germany: Some Stylized Facts About Its Time Path, Causes and Consequences', *Zeitschrift für Wirtschafts- und Sozialwissenschaften* 34, 185–220.

Berger, H. and S. Danninger (2006), 'Does Excessive Regulation Impede Growth In Germany?', *IMF Country Report* 06/17, 106–133.

Berger, H. and M. Neugart (2008), 'Labor Courts, Nomination Bias, and Unemployment in Germany', Manuscript, November.

Bertola, G., T. Boeri and S. Cazes (1999), 'Employment Protection and Labor Market Adjustment in OECD Countries: Evolving Institutions and Variable Enforcement', ILO Employment and Training Papers, 49.

Boyd, C. L., L. Epstein and A. D. Martin (2009), 'Untangling the Causal Effects of Sex on Judging', Manuscript, Northwestern University.

Brudney, J. J., S. Schiavoni and D. J. Merrit (1999), 'Judicial Hostility Toward Labor Unions? Applying the Social Background Model to a Celebrated Concern', *Ohio State Law Journal* 60, 1675–1766.

Däubler, W., J. P. Hjort and D. Hummel (2009), *Arbeitsrecht, Kommentar*, Nomos Verlag, Baden-Baden.

Davis, S., S. Haire and D. R. Songer (1993), 'Voting Behavior and Gender on the U.S. Courts of Appeals', *Judicature* 77, 129–133.

Franz, W. (1994), 'Chancen und Risiken einer Flexibilisierung des Arbeitsrechts aus ökonomischer Sicht', *Zeitschrift für Arbeitsrecht* 25, 439–462.

Galdón-Sánchez, J. E. and M. Güell (2003), 'Dismissal Conflicts and Unemployment', *European Economic Review* 47, 323–335.

Giles, M. W., V. A. Hettinger and T. Peppers (2001), 'Picking Federal Judges: A Note on Policy and Partisan Selection Agendas', *Political Research Quarterly* 54, 623–641.

Goerke, L. (2002), 'On Dismissal Pay', *Labour Economics* 9, 497–512.

Goerke, L. and M. Pannenberg (2009), 'The Effects of Income Taxation on Severance Pay', *Labour Economics* 16, 107–118.

Goerke, L. and M. Pannenberg (2010), 'An Economic Analysis of Dismissal Legislation: Determinants of Severance Pay in West Germany', *International Review of Law and Economics* 30, 71–85.

Grund, C. (2006), 'Severance Payments for Dismissed Employees in Germany', *European Journal of Law and Economics* 22, 49–71.

Hanssen, F. A. (2004), 'Is There a Politically Optimal Level of Judicial Independence?', *American Economic Review* 94, 712–729.

Heckman, J. J. (1976), 'The Common Structure of Statistical Models of Truncation, Sample Selection, and Limited Dependent Variables and a Simple Estimator for Such Models', *Annals of Economic and Social Measurement* 5, 475–492.

Hefeker, C. and M. Neugart (2010), 'Labor Market Regulation and the Legal System', *International Review of Law and Economics* 30, 218–225.

Höland, A., U. Kahl and N. Zeibig (2007), Kündigungspraxis und Kündigungsschutz im Arbeitsverhältnis, Eine empirische Praxisuntersuchung aus Sicht des arbeitsgerichtlichen Verfahrens. 1. Auflage. Nomos, Baden-Baden.

Hromadka, W. and F. Maschmann (2002), Arbeitsrecht Band 1, Individualarbeitsrecht. 2. Auflage. Springer-Verlag, Berlin.

Huang, C., H. Chang and C. Lai (2009), 'Employment Effect of Dismissal Pay in the Presence of Judicial Mistakes', *International Review of Law and Economics* 29, 38–45.

Ichino, A., M. Polo and E. Rettore (2003), 'Are Judges Biased By Labor Market Conditions?', *European Economic Review* 47, 913–944.

Jahn, E. (2005), 'Wie wirkt der Kündigungsschutz', *Zeitschrift für ArbeitsmarktForschung* 2–3, 284–304.

Jahn, E. (2009), 'Do Firms Obey the Law When They Fire Workers? Social Criteria and Severance Payments in Germany', *International Journal of Manpower* 30, 672–691.

Jahn, E. and C. Schnabel (2003), Kündigungsschutz: Reform in Trippelschritten oder besser im großen Sprung, IAB-Kurzbericht Nr. 21 vom 12.12.2003. Available at: http://doku.iab.de/kurzber/2003/kb2103.pdf.

M. Kittner and B. Zwanziger (eds.) (2001), *Arbeitsrecht*, Handbuch für die Praxis, Bund-Verlag.

Norton, E. C., H. Wang and C. Ai (2004), 'Computing Interaction Effects and Standard Errors in Logit and Probit Models', *Stata Journal* 4, 154–167.

OECD (2004), *The OECD Jobs Study*, OECD, Paris.

Peresie, J.L (2005), 'Female Judges Matter: Gender and Collegial Decision Making in the Federal Appellate Courts', *Yale Law Journal* 114, 1759–1790.

Pfarr, H., K. Ullmann, M. Bradtke, J. Schneider, M. Kimmich and S. Bothfeld (2005), *Der Kündigungsschutz zwischen Wahrnehmung und Wirklichkeit: Betriebliche Erfah-rungen mit der Beendigung von Arbeitsverhältnissen*, Rainer Hampp Verlag, München.

Pissarides, C. (2000), *Equilibrium Unemployment Theory*, MIT Press, Cambridge, MA.

Segal, J. A. (2000), 'Representative Decision Making on the Federal Bench: Clinton's District Court Appointees', *Political Research Quarterly* 53, 137–150.

Shapiro, C. and J. M. Stiglitz (1984), 'Equilibrium Unemployment as a Worker Discipline Device', *American Economic Review* 74, 433–444.

Stähler, N. (2008), 'Firing Costs, Severance Payments, Judicial Mistakes, and Unemployment', *Labour Economics* 15, 1162–1178.

Sunstein, C. R., D. Schkade and L. Ellman (2004), 'Ideological Voting on Federal Courts of Appeals: A Preliminary Investigation', *Virginia Law Review* 90, 301–354.

Wooldridge, J. M. (2006), *Introductory Econometrics, A Modern Approach*, 3rd edn, Mason, Ohio: Thomson/South-Western.