Should Non-Expert Courts Control Expert Administrations?

Georg von Wangenheim
Max-Planck-Institute for Research into Economic Systems
Kahlaische Straße 10, 07745 Jena, Germany
wangenheim@mpiew-jena.mpg.de

Paper to be presented at the
19th Annual Meeting of the European Association of Law and Economics
19-21 September 2002, Athens, Greece

September 16, 2002

Abstract

Decisions of public administrations are often disputed in courts. Courts usually have much less expertise in the field of the decision than the administrative body whose original decision is challenged. Therefore the question arises whether administrative decisions which are handed down by experts in their field should actually be subject to court control. In this paper, I first show that control by experts may be better than no control at all, if review is triggered by an appeals process or if review increases administrators' incentives to hand down correct decisions. The simple model also shows that without the incentive effect, the error probability of reviewing courts should be lower than the administration's error probability. In an extension of the model, I discuss the reasons and effect of discretionary ranges for administrative decisions.

1 Introduction

Decisions of public administrations are often disputed in courts. Courts usually have much less expertise in the field of the decision than the administrative body whose original decision is challenged. This difference in expertise will often imply that courts hand down incorrect decisions with a higher probability than administrators. Therefore the question arises whether administrative decisions which are handed down by experts in their field should actually be subject to court control.

Court control has two central aspects which may improve administrative decision making even if courts lack expertise and therefore are more likely to err than administrators. One aspect is that court review usually is triggered by parties who are negatively affected by the administrative decision and who have some private information on the specific case. As SHAVELL (1995) shows for appeals of court decisions, the costs which an appelant has to bear may (partly) separate appelants whose defeat in the administrative procedure was erroneous from those who lost with good reasons.

The other aspect is that court review may induce administrators who dislike being reversed to exert more error avoiding effort. While this aspect fails to select a wrong administrative decisions for judicial review with a higher probability than correct administrative decisions, it improves the quality of administrative decisions directly. Administrators have more incentives to decide correctly and will thus exert more effort on avoiding errors.[1]

This paper shows how these to aspects of judicial review work. In particular, I will show for a simple model based on SHAVELL's paper with two possible types of projects — one welfare increasing and one welfare reducing — that the the appeals process even with the non-expert administrative courts reduces the probability of incorrect decisions on projects. Similarly, random review of administrative decisions by non-expert courts may reduce total error probabilities. Thus both aspects alone may justifiy the costs of resources necessary to install a judicial review process by non-expert courts.

However, the model is not restricted to the question whether non-expert control may justify its costs. In addition, it also allows to inquire whether

[1] Note that these two control approaches correspond closely to the question of whether *political* control of administrations should be done in a fire alarm manner or like a police patrol, cf. MCCUBBINS AND SCHWARTZ (1984) and WEINGAST (1984) for informal treatments and BAWN (1997) for a formal model; VON WANGENHEIM (2002: section 3) gives an overview.

control should be performed by non-experts with higher error probabilities than the administrators or rather by experts with lower error probabilities than the administrators.

Shavell's model with only two alternative project types suffers from a severe shortcoming: a Nash equilibrium in pure strategies only exists if courts fail to infer any information from their knowledge of who filed the appeal. Otherwise the unique Nash equilibrium is on in mixed strategies, whose predictive power for real human behavior has been challenged with good reasons (cf. e.g. Samuelson (1997: 27)). I therefore extend the model to account for a two-dimensional continuum of possible projects: both the gains of the owner of a project and the harm inflicted on third parties may range from zero to infinity. With such a modelling approach, it turns out that some appelants will appeal administrative decisions which are welfare improving. The question whether the appeals process necessarily increases social welfare if one leaves the resource costs of judicial review aside therefore arises anew. As a side product of this discussion, it will turn out that the model also provides a theory on judicial deference to administrative decision making. This theory explains why benevolent courts sometimes follow this doctrine and somtimes do not.

Before presenting the arguments of the paper, some definitions nedd clarification. Troughout the paper, I will assume that the aim of administrative law is to improve social welfare, which is defined as the difference between gains and harm resulting from all those projects which are subject to regulation by administrative law. I thus use the expressions "legal", "correct", and "welfare improving" as synonyms. Accordingly, there is also no difference between the terms "illegal", "wrong", and "welfare reducing". I also assume that all gains and harm resulting from projects are well defined and and may be obesrved by some individuals at least after the project has been carried out. Further, I use the term "expertise" to denote the ability to seperate welfare improving from welfare reducing projects. Hence an expert is more likely to to find out correctly whether a specific project increases or reduces welfare than a non-expert. Expertise thus does not refer to superior knowledge on a subset of the qualities which influence the welfare effects of projects, but to the entirety of these qualities. FInally, the "owner" of a project is the indicvidual who can carry out the project from a factual point of view and who will gain directly from the realization of the project. He may become an "applicant" if he files for a public law permission for his project with the competent authorities and an "appelant" if he appeals the rejection of his application by the administration. Opposed to the owner of the project are third parties who will only bear the costs (harm) from realization of the project. They can also become "appelants" if they appeal the administrative approval of the project.

The paper starts with the discussion of the welfare effects of non-expert control as compared to no control (section 2). I will first present the central argument in an informal way (section 2.1) and then underpin the argument in a technical way (sections 2.2 through 2.4). Section 3 compares non-expert control to expert control and section 4 extends the model to a two-dimensional continuum of possible projects. After presenting the structure of the model (section 4.1), I again present the central arguments both in an informal way (section 4.2) and in a formal way (section 4.3). Section 5 concludes.

2 Poor control versus no control

2.1 Informal argument

In most modern democracies, administrative decisions are subject tot judicial control. Frequently, the reviewing courts appear to be more err more often on the welfare effects of their decisions than the reviewed administration. I will show in the following, that judicial control may nevertheless increase social welfare for two reasons: because it is triggered by appeals of losing parties and because it adds to the incentives of administrators to hand down correct decisions.

The idea for the appeals process is the following. Parties who are negatively affected by administrative decisions decide on appealing the decision by comparing their expected gains to the cost of appealing. While the costs tend to be independent of the merits of an appeal, the expected gains increase in the legal merits as long as courts reverse the administrative decisions with higher probabilities when the legal merits of the appeal are higher. As a consequence, wrong administrative decisions are more likely to be appealed than correct ones. Thus, the unappealed decisions are on average more likely to be correct than the appealed decisions and thus than all administrative decisions together. As long as the courts' error probabilities are lower than the proportion of incorrectly decided cases among the appealed cases, the appeals process reduces the overall error probability. The courts' error probability thus need not be below the administrations' error probability. The self-selection of appelants compensates for the difference. Depending on its costs, the appeals triggered review may thus be justified by a reduction of overall error probabilities even if courts are more error prone than administrators.

If review is not triggered by appeals but cases are randomly selected for review but administrators dislike to see their decisions being reversed by courts, again the error probability of the unappealed administrative decisions decreases. This time not because correct decisions are more likely to remain unappealed but because administrators work harder on avoiding errors. Once more, this may offset poor decision making by the courts. Now the courts' error probabilities may even be higher than the administration's error probability among the reviewed cases. However, the reduced error probabilities of the administration must offset both the large error probabilities of the courts and their resource costs. Obviously, the reduction in administrative error probabilities becomes small if administrators hardly can influence the reversal probability because court decisions are nearly independent of the legal merits of the reviewed cases.

In the following, I will refine these intuitive arguments in a formal model based on Shavell (1995).

2.2 Formal argument: appeals process

Let there be two types of projects, one welfare increasing and legal and one welfare decreasing and illegal. The welfare effects of the projects are $w_h > 0$ for the legal project and $w_\ell < 0$ for the illegal one. Let the proportion of legal plans among the applications be given by f^o. Administrators have incomplete information about the legal merits of the projects brought before them. They investigate the cases to find out each project's true type, but they err with probability $q_A \in (0, 1/2)$. Judges decide similar to administrators: they investigate the cases which are appealed to find out each project's true type and err with probability $q_J \in (q_A, 1/2)$, i.e. their error probability is larger than the administrators' error probability. This reflects their lacking expertise.

After the administration has handed down its decision on a project, both the beneficiaries of the project — typically one applicant who wants to carry out the project but needs a public law permit to do so — and those (third) parties who suffer from the project may appeal the administration's decision if it is to their disadvantage. The costs of filing an appeal for applicants and third parties are $c_1 > 0$ and $c_3 > 0$, where the index denotes first and third party, respectively. These costs consist of the resources the appelant spends on preparing and defending the appeal and possibly a state imposed fee or subsidy. The gains from a successful appeal are $g_x = w_x + h_x > w_x$ for $x \in \{h, \ell\}$ for the applicant and $\lambda h_x = \lambda(g_x - w_x) > -\lambda w_x$ for $x \in \{h, \ell\}$ for the the third party which is most affected and bears a fraction $\lambda \in (0, 1)$ of the total harm.[2] applicants and third parties are completely informed about the values of g and h of the relevant project. Aministrators and courts only know which values are possible and the probability f^o.

Applicants appeal a disfavorable decision if and only if $c_1 < g_h(1 - q_J)$ if the project is legal and $c_1 < g_\ell q_J$ if the project is illegal. Similarly, third parties appeal a project approving decision if and only if $c_3 < \lambda h_\ell(1 - q_J)$ if the project is illegal and $c_3 < \lambda h_h q_J$ if the project is legal.

Following Shavell, I call the appeals process "perfectly separating" if all and only incorrect decisions of the administration are appealed. This is the case if

$$c_1 \in (g_\ell q_J, g_h(1 - q_J)) \quad \text{and} \quad c_3 \in (\lambda h_h q_J, \lambda h_\ell(1 - q_J)). \tag{1}$$

Note that at least one of the intervals must exist due to $1 - q_J > q_J$ and $w_\ell = g_\ell - h_\ell < w_h = g_h - h_h$ which implies $h_h - h_\ell < g_h - g_\ell$ so that either $g_h > g_\ell$ or $h_\ell > h_h$ or both. For the sake of the argument, assume that both intervals exist and that the appeals costs satisfy both parts of (1).

Measure social welfare by the expected welfare effect per application

$$W = f^o p_h w_h + (1 - f^o) p_\ell w_\ell - p_{\text{appeal}} c_J \tag{2}$$

where p_h and p_ℓ are the probabilities that a project of type h or ℓ, respectively, passes both the administration and a possible court review, p_{appeal} is the probability that an appeal is lodged independently of the administrative decision and the legal merits of the project, and c_J are the social costs of the resources spent on each appeal. For the perfectly separating appeals process, i.e. if all and only administrative errors are appealed, social welfare then is

$$W_{\text{perfect}} = f^o(1 - q_A q_J) w_h + (1 - f^o) q_A q_J w_\ell - q_A c_J. \tag{3}$$

[2]Keeping λ fixed is a crude simplification. I conjecture that this simplification does not substantially affect the results.

Given that without any review, social welfare is $W_o = f^o(1-q_A)w_h + (1-f^o)q_A w_\ell$, the perfectly separating appeals process is socially superior to the absence of any appeal if

$$q_J < 1 - \frac{c_J}{f^o w_h - (1-f^o)w_\ell}. \tag{4}$$

Due to $f^o w_h - (1-f^o)w_\ell > 0$, this is compatible with at least some $q_J \in (q_A, 1/2)$, if the social appeals costs c_J are sufficiently small. Hence, even if courts have a higher error probability than the administration, control may be worthwhile if it is based on a perfectly separating appeals process. Restating the above condition as

$$c_J < (1-q_J)(f^o w_h - (1-f^o)w_\ell) \tag{5}$$

makes it more intuitive: the (social) costs of each appeal must be less than the expected gain from the appeal, i.e. the probability that the incorrect administrative decision is repealed times the expected loss accruing from an incorrect administrative decision.

2.3 Formal argument: incentive effects

To isolate the incentives effect of judicial review, I assume in this section, that court review is not triggered by appeals but by random selection of a subset of the administrative decisions. While I stick to the assumption that judges err on the legal merits of a project with a fixed probability $q_J \in (q_{A0}, 1/2]$ where q_{A0} is the administrative error probability without court review, I now assume that the administrators' probability to misjudge the project depends on the probability that courts will reverse their decision.

In particular, I assume that each administrator's expected utility deciding on n cases is the sum of some intrinsic motivation to decide correctly, an "extrinsic" motivation to avoid court reversals,[3] and the negative of his disutility of effort:

$$\begin{aligned} \mathrm{E}U_{\text{stoch}} &= -nrs_x(q_A(e_A)(1-q_J) + (1-q_A(e_A))q_J) \\ &\quad + ns_i\left(f^o(1-q_A(e_A))w_h + (1-f^o)q_A(e_A)w_\ell\right) - u(ne_A) \end{aligned} \tag{6}$$

[3]The "extrinsic" motivation may be due to sanctions e.g. of superiors in the form of promotion probabilities or of social sanctions from peers but may also be due to a simple dislike of being reversed, i.e. a sanction which is not extrinsic in a strict sense.

where r is the review probability, s_x is the sanction describing the "extrinsic" motivation, s_i is the sanction describing the "intrinsic" motivation, $q_A(e_A)$ is the the administrator's error probability as a function of the effort e_A he exerts per case with the properties $q_A(0) \leq 1/2$, $\lim_{e_A \to \infty} = 0$, $q'_A < 0$, and $q''_A > 0$, and $u(ne_A)$ is the disutility of effort. The parameter n is included only in order to kep the argument consistent with a similar extension for the judges to be introduced in section 3. For the time being, however, I will assume that the courts' error probability remains exogenously given as q_J.

Taking the derivative of the expected utility $\mathrm{E}U_{\text{stoch}}$ with respect to the administrator's effort e_A and setting the result equal to zero

$$\frac{\mathrm{dE}U_{\text{stoch}}}{\mathrm{d}e_A} = -n\Big(s_i(f^o w_h - (1-f^o)w_\ell) + rs_x(1-2q_J)\Big)q'_A(e_A) - nu'(ne_A) \overset{!}{=} 0 \tag{7}$$

yields the administrator's optimal effort e^*_A as a function of $r \cdot s_x$ with

$$\frac{\mathrm{d}e^*_A(rs_x)}{\mathrm{d}(rs_x)} = -rs_x \frac{q''_A(e^*_A)}{q'_A(e^*_A)} + \frac{u''(ne^*_A) + s_i(f^o w_h - (1-f^o)w_\ell)q''_A(e^*_A)}{-(1-2q_J)q'_A(e^*_A)} \tag{8}$$

which is strictly positive. One should note that the incentive effect only depends on the product rs_x. As a consequence, with sufficiently large sanctions s_x, an arbitrarily small r may induce every effort level and thus every error probability. Of course, all caveats against high sanctions with low enforcement probabilities in criminal law and economics[4] apply here as well. In addition, the participation constraint of the administrator may become relevant and induce additional social costs of inducing high efforts. As long as the administrator's utility is linear in the extrinsic sanction and his wage, a flat increase of wage may offset his loss in expected utility without increasing the costs to society. Only if his utility is non-linear in the extrinsic motivation or in his wage, satisfying the administrator's participation constraint (i.e. keeping him on the job) levies aditional costs on society of the expected extrinsic sanctions increase. Of course, if the participation constraint is not binding, i.e. if the administrator extracts rents from his job, there are no such costs to society.

Neglecting these problems, one can show that a purely stochastic review procedure with a sufficiently small review probability r and a sufficiently large

[4] Cf. XXX for an overview.

external sanction s_x is superior to the absence of control of administrative decisions even with the rather strong assumption $q_J \in (q_{A0}, 1/2]$ where $q_{A0} = q_A(e_A^*(0))$.

Social welfare with stochastic review is given by

$$\begin{aligned} W_{\text{stoch}} &= (1-r)\Big(f^o(1-q_A(e_A^*(rs_X)))w_h + (1-f^o)q_A(e_A^*(rs_X))w_\ell\Big) \\ &\quad +r\Big(f^o(1-q_J)w_h + (1-f^o)q_J w_\ell - c_J\Big). \end{aligned} \tag{9}$$

Hence, stochastic review is socially superior to the absence of any review of administrative decisions if and only if

$$q_J < \frac{q_A(e_A^*(0)) - q_A(e_A^*(rs_x))}{r} + q_A(e_A^*(rs_x)) - \frac{c_J}{f^o w_h - (1-f^o)w_\ell}. \tag{10}$$

Whether the first term on the right-hand side of this equation increases or decreases in r, depends on the exact shape of the functions $q_A(\cdot)$ and $u(\cdot)$ and on the level of s_x. However, if one keeps the product $rs_x > 0$ constant, one can increase this term to any desired level by reducing the review probability r. Thus, if it is possible to costlessly increase the external sanction s_x, every courts' error probability q_J is small enough to make stochastic review with a sufficiently low review probability superior to no control of administrative decisions.

Only if one cannot increasse s_x in a costless way, the judicial error probability may be too large for the superiority of stochastic review. Of course, this result becomes more likely the larger the resources c_J necessary to perform court control.

2.4 Formal argument: combining appeals and incentives effects

If one includes the incentives effect in a model of perfect appeals, the resulting social welfare effect will be strictly larger than without the incentive effect. However, replacing random selection of cases for review by an appeals process may reduce social welfare, if this implies an increase in the number of cases reviewed.

With a perfect appeals process and the incentives effect, the administrators' expected utility becomes

$$\begin{aligned} \mathrm{E}U_{\text{p\&i}} &= -ns_x q_A(e_A)(1-q_J) \\ &\quad +ns_i\left(f^o(1-q_A(e_A))w_h + (1-f^o)q_A(e_A)w_\ell\right) - u(ne_A) \end{aligned} \tag{11}$$

since the review probability of correct administrative decisions is zero and the review probability of incorrect decisions is one. The administrators' optimality condition then becomes

$$\frac{\mathrm{dE}U_{\text{p\&i}}}{\mathrm{d}e_A} = -n\Big(s_i(f^o w_h - (1-f^o)w_\ell) + s_x(1-q_J)\Big)q'_A(e_A) - nu'(ne_A) \overset{!}{=} 0$$

As one can easily see, the resulting optimal effort $e_A^{**}(s_x)$ of the administrator increases in s_x and is strictly larger than with random selection of cases for review since the cofactor of $q'_A(e_A)$ becomes larger in absolute terms due to $s_x(1-q_J) > rs_x(1-2q_J)$.

Social welfare with a perfectly separating appeals process and a positive incentives effect ($W_{\text{p\&i}}$) is the same as in equation (3) with q_A replaced by $q_A(e_A^{**}(s_x)$. Obviously, social welfare increases in s_x since $q'_A(\cdot) < 0$ and $\frac{\mathrm{d}W_{\text{p\&i}}}{q_A} < 0$. Hence, the social welfare effect of appeals triggered review is larger if an incentive effect exists than if it does not.

However, social welfare with appeals triggered review is unambiguously larger than with stochastic review *only* if the number of appeals is not larger than the number of random reviews (i.e. $q_A(e_A^{**}(s_x)) \leq r$). Otherwise, the second term of the difference

$$\begin{aligned} W_{\text{p\&i}} - W_{\text{stoch}} &= (1-r)q_A(e_A^*(rs_x))(f^o w_h - (1-f^o)w_\ell) \\ &\quad + \Big(r - q_A(e_A^{**}(s_x))\Big)\Big(c_J + q_J(f^o w_h - (1-f^o)w_\ell)\Big) \end{aligned} \tag{12}$$

is negative and may outweigh the first term. A numerical example shows that the difference may become negative. With the values

$$\left.\begin{array}{c} r = .01, \quad f^o = .5, \quad w_h = 1, \quad w_\ell = -1, \quad q_J = .29, \quad c_J = .8 \\[2ex] q_A(e_A^{**}(s_x)) = .26, \quad q_A(e_A^*(rs_x)) = .27, \quad q_A(e_A^*(0) = .28 \end{array}\right\} \tag{13}$$

one gets

$$W_{\text{stoch}} = 0.2218 > W_o = 0.22 > W_{\text{p\&i}} = 0.2166 > W_{\text{perfect}} = 0.1948. \tag{14}$$

Here, stochastic review is the optimal solution while perfectly separating appeals as a trigger for review is the worst approach. Of course, perfectly separating appeals as a trigger for review with incentives effects is better, but still worse than no review at all.

The intuition is straightforward. With stochastic review of one percent of all decisions, one can induce a reduction of the error probability of one per cent for all other decisions at the cost of one per cent additional error probability for the one percent of reviewed cases. The effect on the average error probability is thus a reduction by 0.98 per cent. For this reduction, society has to incurr the resource costs of reviews of one per cent of all cases. With appeals as a trigger for reviews instead of fandom selection, all 26 per cent of false decisions are reviewed, the resources society spends on review are thus 26 times as high. The error probability of the entire process, however, declines only from $(1-r)q_A(e_A^*(rs_x)) + rq_J = 0.2702$ to $q_A(e_A^{**}(s_x))q_J = 0.0754$, i.e. by 0.1948. This reduction is less than 26 times the error reduction of 0.0098 induced by the transition from no review to stochastic review. Hence, the resources which have to be spent for court reviews may be justifed by stochastic review with a small review probability but not by appeals triggered review.

One should not however, that the result changes, if one leaves the number of reviews unchanged, i.e. if one assumes that not all false decisions of the administration are reviewed but only a proportion thereof which results in the same total number of reviews as under stochastic review. Then the administrators' expected utility becomes

$$\begin{aligned} \mathrm{E}U_{\mathrm{p\&i\&r_a}} &= -ns_x r_a q_A(e_A)(1-q_J) \\ &\quad +ns_i\left(f^o(1-q_A(e_A))w_h + (1-f^o)q_A(e_A)w_\ell\right) - u(ne_A) \qquad (15) \end{aligned}$$

where $r_a = r/q_A(e_A^*(rs_x))$ is the review probability of stochastic review divided by the error probability of the administrator, i.e. the probability that appealed decisions are reviewed. Note that if the administrator includes this dependence into his optimization, then he knows that e.g. one per cent of his decisions will be reviewed and that all these decisions are false (assuming that he will not be able to achieve a lower error probability with reasonable effort), and hence that the amount of sanctions for his being reversed by the courts is independent of his effort level. He will thus exert as little effort as without any extrinisic incentives effect. If, however, he takes the review probability of his false decisions as given and independent of his own error probability, then his optimization condition becomes

$$\begin{aligned} &\frac{\mathrm{dE}U_{\mathrm{p\&i\&r_a}}}{\mathrm{d}e_A} = \\ &\quad -n\Big(s_i(f^o w_h - (1-f^o)w_\ell) + r_a s_x(1-q_J)\Big)q_A'(e_A) - nu'(ne_A) \stackrel{!}{=} 0 \end{aligned}$$

Comparing this conditions to the earlier optimality conditions shows that the resulting effort $e_A^{***}(s_x r_a)$ is less than the effort with review of all appeals in a perfectly separating appeals process $e_A^{**}(s_x)$ but more than the effort with strictly stochastic review $e_A^*(rs_x)$.

Social welfare with this limited appeals triggered review is

$$\begin{aligned} W_{\mathrm{p\&i\&r_a}} &= f^o(1-q_A(1-r_a(1-q_J)))w_h \\ &\quad +(1-f^o)q_A(1-r_a(1-q_J))w_\ell - q_A r_A c_J. \end{aligned}$$

Since $r_A = r/q_A(e_A^*(rs_x))$ by definition, the difference

$$\begin{aligned} W_{\mathrm{p\&i\&r_a}} - W_{\mathrm{stoch}} &= (r(1-q_A(e_A^*(rs_x))) + q_A(e_A^*(rs_x)) - q_A(e_A^{***}(s_x r_a))) \\ &\quad \cdot(f^o w_h - (1-f^o)w_\ell) \end{aligned}$$

is strictly positive due to $q_A(e_A^*(rs_x)) > q_A(e_A^{***}(s_x r_a))$. The intuition is obvious: with an appeals based review under which the same number of cases are reviewed as under the stochastic review, the courts cost the same amount of resources, but may only correct false decisions or leave them unchanged while under the stochastic review process they may also reverse correct decisions.

Thus appeals processes with perfectly separating appeals but only partial judicial review of appealed decisions may be the optimal decision.

3 Poor Control versus Good Control

In the previous section I have shown that costly control of administrative decisions by courts may be justified even if courts commit more errors than the administration. The result, was based, however, on fixed error probabilities of the courts and fixed resources spent on administrative decision making. In this section, I will take the resources spent on administrative and judicial decision making as policy variables which induce different effort levels of administrators and judges. The background of this induction is simple: the more resources society spends on decision making, the more decision makers may be employed and thus the less cases each decision maker has to decide upon. This reduces his marginal disutility of effort for each single case and thus results in higher effort per case. In the following, I will show that resources should be allocated so that courts have a lower error probability than administrators if non-expert control of administrative decisions can only be justified by the appeals process and not by the incentives effect of review (cf. section 2.2). In other words, non-expert control may be better than no control, but expert control is better than non-expert control. Only if one adds the incentives effect, the result may change, though only in extreme cases. A variation of the assumptions allowing judges to be non-experts in the sense that they need more effort to achieve the same error probability as administrators, will also change the results: the optimal error probability of reviewing courts may be larger than the optimal error probability of administrators.

To formalize the argument, I first assume that there is a given large pool of decision makers which may either serve as administrators or as judges. Their abilities to avoid errors given a certain amount of effort do not differ whether they work as administrators or as judges. Employment of each decision maker costs an amuont c to society. With a perfectly separating appeals process as trigger for judicial review, total social welfare is then given by

$$W_n = n(f^o w_h(1 - q_A(e_A)q_J(e_J)) + (1 - f^o)w_\ell q_A(e_A)q_J(e_J)) - c(M_A + M_J)$$

where n is the total number of cases (not per administrator as before but now for the entire society), e_A is the effort exerted per case by a decision maker working as administrator (as before) and e_J is the effort exerted per case by a decision maker working as judge. M_A and M_J are the numbers of decision makers employed as administrators or judges, respectively.

To determine how the effort levels of the decision makers depend on the numbers of decision makers, I consider their expected utilities. As before, I assume that administrators are benevolent to some degree in the sense that they have an intrinsic motivation to hand down correct decisions.[5] Their disutility of effort continues to depend on total effort exerted and they may have an extrinsic motivation to avoid court revesals of their decisions. Their expected utility is thus given by

$$\begin{aligned} \mathrm{E}U_n = & -\frac{n}{M_A} s_x q_A(e_A)(1 - q_J(e_J)) \\ & + \frac{n}{M_A} s_i \left(f^o(1 - q_A(e_A))w_h + (1 - f^o)q_A(e_A)w_\ell\right) - u(e_A \frac{n}{M_A}) \end{aligned} \tag{16}$$

where n/M_A is the number of cases on which the single administrator has to decide and the other variables remain defined as before.

[5] Again, I assume for simplicity that they do not adjust their benevolence by the court corrections of their decisions.

For the judge, expected utility has the same structure except for the extrinsic motivation which I assume not to exist for judges. This is counterfactual if court decisions are also subject to review. However, adding more court levels would increase the complexity of the argument and not provide additional insight since the structure of the argument would remain the same. Only if one investigates the optimal number of appeals levels, one has to make these levels explicit. This extension of the argument has to be left to further research. Assuming that no extrinsic motivation for judges exists thus amounts to reducing court review to just one court level. Then a judge's expected utility is given by

$$\mathrm{EV}_n = \frac{m}{M_J} s_i \left(f^o(1-q_J(e_J))w_h + (1-f^o)q_J(e_J)w_\ell\right) - u(e_J\frac{m}{M_J}) \tag{17}$$

where $m = nq_A(e_A)$ is the total number of cases under review and m/M_J is the number of cases under review per judge. Note that the proportions of legal and illegal projects in the set of cases under judicial review is the same as in the set of all cases because the error probability of administrators does not differ for legal and illegal projects.

Both administrators and judges optimize their behavior by equating the first derivatives to zero:

$$\frac{\mathrm{dEU}_n}{\mathrm{d}e_A} = -\frac{n}{M_A}\Big(\big(s_i\tilde{w}+s_x(1-q_J(e_J))\big)q_A'(e_A)+u'(e_A\frac{n}{M_A})\Big) \stackrel{!}{=} 0 \tag{18}$$

and

$$\frac{\mathrm{dEV}_n}{\mathrm{d}e_J} = -\frac{nq_A(e_A)}{M_J}\Big(s_i\tilde{w}\,q_J'(e_J) + u'(e_J\frac{nq_A(e_A)}{M_J})\Big) \stackrel{!}{=} 0 \tag{19}$$

where I use the abbreviation $\tilde{w} \equiv f^o w_h - (1-f^o)w_\ell$. This simultaneous equation system defines the decision makers' optimal values of e_A and e_J as functions of the policy variables M_A and M_J. Taking total derivatives of both functions and applying Cramer's rule, one gets

$$\frac{\mathrm{d}e_A}{M_K} = -\begin{vmatrix} \frac{\partial^2 \mathrm{EU}_n}{\partial e_A \partial M_K} & \frac{\partial^2 \mathrm{EU}_n}{\partial e_A \partial e_J} \\ \frac{\partial^2 \mathrm{EV}_n}{\partial e_J \partial M_K} & \frac{\partial^2 \mathrm{EV}_n}{\partial e_J^2} \end{vmatrix} \Bigg/ \begin{vmatrix} \frac{\partial^2 \mathrm{EU}_n}{\partial e_A^2} & \frac{\partial^2 \mathrm{EU}_n}{\partial e_A \partial e_J} \\ \frac{\partial^2 \mathrm{EV}_n}{\partial e_A \partial e_J} & \frac{\partial^2 \mathrm{EV}_n}{\partial e_J^2} \end{vmatrix} \qquad K \in \{A, J\} \tag{20}$$

for the effect of M_A and M_J on e_A and

$$\frac{\mathrm{d}e_J}{M_K} = -\begin{vmatrix} \frac{\partial^2 \mathrm{EU}_n}{\partial e_A^2} & \frac{\partial^2 \mathrm{EU}_n}{\partial e_A \partial M_K} \\ \frac{\partial^2 \mathrm{EV}_n}{\partial e_A \partial e_J} & \frac{\partial^2 \mathrm{EV}_n}{\partial e_J \partial M_K} \end{vmatrix} \Bigg/ \begin{vmatrix} \frac{\partial^2 \mathrm{EU}_n}{\partial e_A^2} & \frac{\partial^2 \mathrm{EU}_n}{\partial e_A \partial e_J} \\ \frac{\partial^2 \mathrm{EV}_n}{\partial e_A \partial e_J} & \frac{\partial^2 \mathrm{EV}_n}{\partial e_J^2} \end{vmatrix} \qquad K \in \{A, J\} \tag{21}$$

for the effect of M_A and M_J on e_J. Considering the signs of the derivatives

$$\begin{aligned}
\frac{\partial^2 \mathrm{E}U_n}{\partial e_A^2} &= -\frac{n}{M_A}\left(\left(s_i\tilde{w}+s_x(1-q_J(e_J))\right)q_A''(e_A)+u''\!\left(e_A\frac{n}{M_A}\right)\frac{n}{M_A}\right)<0\\
\frac{\partial^2 \mathrm{E}U_n}{\partial e_A \partial e_J} &= \frac{n}{M_A}\, s_x\, q_J'(e_J)\, q_A'(e_A)\\
\frac{\partial^2 \mathrm{E}V_n}{\partial e_A \partial e_J} &= -\frac{n\, q_A(e_A)}{M_J}\, u''\!\left(e_J\frac{n\, q_A(e_A)}{M_J}\right) e_J\, \frac{n\, q_A'(e_A)}{M_J}>0\\
\frac{\partial^2 \mathrm{E}V_n}{\partial e_J^2} &= -\frac{n\, q_A(e_A)}{M_J}\left(s_i\tilde{w}\, q_J''(e_J)+u''\!\left(e_J\frac{n\, q_A(e_A)}{M_J}\right)\frac{n\, q_A(e_A)}{M_J}\right)<0
\end{aligned}$$

one can easily see that the denominator of the expressions is positive unless the external incentives (s_x) are very large. Only then the product $\frac{\partial^2 \mathrm{E}U_n}{\partial e_A \partial e_J}\frac{\partial^2 \mathrm{E}V_n}{\partial e_A \partial e_J}$ is larger than $\frac{\partial^2 \mathrm{E}U_n}{\partial e_A^2}\frac{\partial^2 \mathrm{E}V_n}{\partial e_J^2}$ so that the denominator becomes negative. However, if it does, the slope of the function defined by $\frac{\partial EU_n}{\partial e_A}=0$ in the e_A–e_J–space is smaller than the slope of of the function defined by $\frac{\partial EV_n}{\partial e_J}=0$ which implies that the intersection is not a stable equilibrium. I abstract from such extremely large s_x. The numerators (evaluated at the optimization conditions of administrators and judges) are all strictly negative (see appendix A for the exact values). Thus, as long as s_x is not too large, both e_A and e_J increase in both M_A and M_J.

Making use of these derivatives, I will show in the following that at least for no extrinsic incentives of the administrators — and by continuity of the relevant functions also for sufficiently small extrinsic incentives — equal effort levels of administrators and judges cannot be optimal. Rather, one can increase social welfare by increasing the effort of judges above the effort of administrators. This implies that in a social welfare maximum the error probabilities of judges must be lower than the error probabilities of administrators. In this sense, the courts who control the administration should be more expert than the administrators themselves. I will discuss some of the restrictive assumptions needed to derive this result after showing how it emerges.

Rewrite social welfare as

$$\begin{aligned}
W_n = n\Big(& f^o w_h(1-q_A(e_A(M_A,M_J))q_J(e_J(M_A,M_J))) \\
& +(1-f^o)w_\ell q_A(e_A(M_A,M_J))q_J(e_J(M_A,M_J))) - c(M_A+M_J\Big),
\end{aligned} \tag{22}$$

take the derivatives with respect to M_A and M_J, and evaluate at $e_J = e_A$ to get

$$\begin{aligned}\frac{\mathrm{d}W_n}{\mathrm{d}M_A} =& -c - n(f^o w_h - (1-f^o)w_\ell)q_A(e_A(M_A, M_J)) \\ & \cdot q_A'(e_A(M_A, M_J)) \left(\frac{\partial e_A(M_A, M_J)}{\partial M_A} + \frac{\partial e_J(M_A, M_J)}{\partial M_A} \right)\end{aligned} \tag{23}$$

and

$$\begin{aligned}\frac{\mathrm{d}W_n}{\mathrm{d}M_J} =& -c - n(f^o w_h - (1-f^o)w_\ell)q_A(e_A(M_A, M_J)) \\ & \cdot q_A'(e_A(M_A, M_J)) \left(\frac{\partial e_A(M_A, M_J)}{\partial M_J} + \frac{\partial e_J(M_A, M_J)}{\partial M_J} \right).\end{aligned} \tag{24}$$

Now assume $s_x = 0$. Then the optimality conditions of administrators and judges (equations (18) and (19)) together with $e_A = e_J$ imply that $M_J = q_A(e_A)M_A$. Inserting this into the exact values of the derivatives determined in equations (20) and (21) (also cf. appendix A), some basic calculations show that the difference $\frac{\mathrm{d}W_n}{\mathrm{d}M_A} - \frac{\mathrm{d}W_n}{\mathrm{d}M_J}$ reduces to a strictly negative term. Hence absent the incentives effect, $\frac{\mathrm{d}W_n}{\mathrm{d}M_A} < \frac{\mathrm{d}W_n}{\mathrm{d}M_J}$ at $e_A = e_J$. Thus $e_A = e_J$ cannot be an optimum. However, if one increases M_J relative to M_A starting from the values which induced $e_A = e_J$, social welfare will increase. This implies that e_J will become relatively larger than e_A. Hence optimality of effort levels requires $e_J > e_A$ — judges should have lower error probabilities than administrators. Since all functions are continuous, I conjecture that the same result occurs for strictly positive but sufficiently small values of s_x, i.e. for sufficiently small extrinsic motivation of administrators.

The intuition behind this result is rather simple. Due to the assumption of perfect separation of appelants, judges only face the danger of one type of error — false approval of an administrative decision — while administrators face the danger of both error types — false denials of permissions and false granting of permissions. Thus decisions by courts are more likely to increase social welfare than administrative decisions even if both exert the same amount of effort and have the same error avoidance technology.

The argument becomes weaker and possibly reversed if s_x is positive and sufficiently large. It may also be reversed, if one drops the assumption that administrators and judges have the same error avoidance technology. If one instead assumes that effort employed by judges is less productive than effort of administrators since judges have to cover a wider range of different cases and thus cannot specialize in the same way as administrators, then the judges advantage of not being able to commit both types of error may be offset and thus the above result may collapse.

The central insight of this section was that the appeals process alone is not able to justify control of administrators by judges who have a higher error probability than the administrators. Only if the alternative is no control at all, such weak control may be justified on the basis of the appeals process.

4 Continuous Set of Applications

4.1 Structure of the Model

The argument of the previous sections was based on the assumption that courts infer no information from the fact that an appeal was lodged nor from the decision of the administration. Shavell (1995: 393 and 412) justified this simplification by the problems of finding equilibria if courts infer information from the lodging of an appeal. His central argument was that this would imply to reverse all appealed decisions and thus destroy the separating effect of the appeals process. He concludes that procedural rules should forbid courts to infer information from the fact that an appeal was lodged.

However, Shavell's argument is too simple. The structure of the problem is that of an enforcement game. An equilibrium in pure strategies fails to exist. If the courts use the information implied in which of the (fully informed) parties lodges an appeal, then, as Shavell rightly points out, no separation occurs, all parties unhappy with the administration's decision will appeal. However, then the information on who lodged the appeal is worthless, the courts will not use this information but rely on their investigations. With optimal appeals costs, however, this implies that only incorrect administrative decisions are appealed, the courts optimal reply is to rely on the information who lodghed the appeal.

The unique Nash equilibrium of this game is one in mixed strategies. The courts will rely on the information conveyed by the appeal in some cases and disregard this information in the other cases. They will rely on the information conveyed by the appeal in exactly the proportion of cases which makes parties whose case was rightly dismissed before the administration indifferent between lodging an appeal or not (note that this implies that all parties whose case was wrongly dismissed before the administration will appeal). On the other hand, all parties whose case was rightly dismissed before the administration will appeal with a probability which makes the courts indifferent between reliance on prior information (conveyed by the appeal) and investigation of all cases.

It is, however, doubtful whether such mxied strategy Nash equilibria may serve as a reasonable predictor of human behavior. In particular, stability of such equilibria is highly debated (e.g. SAMUELSON (1997: 27)). I therefore refrain from discussing such an equilibrium in any detail. I will rather extend the model to include more than two possible project types. Specifically, I will allow for continua of possible gains and harms from realization of projects, which will result in Nash equilibria in pure strategies. The model then consists of projects which go through the following steps.

Step 0: The project is drawn randomly from a distribution with density $f(g, h)$ on the support R_0^{+2} where g denotes the gains accruing to the owner of the project in case of realization and h denotes the total harm which third parties suffer if the project is realized. The more g and h are correlated, the more equally they grow. In the extreme, if they are perfectly correlated, their difference remains constant.[6] The distribution is common knowledge. The owner of the project knows the realization of g and third parties know the realization of h.

Step 1: The owner of the project may file for permission of his project or dismiss the project. If he files for a permission, he incurs costs of c_o. If he dismisses the project, the game ends. As one may expect and as I will show later, the owner will file for permission if and only if $g \geq \underline{g} \geq c_o$ because only then he may expect his application costs to be covered.[7]

Step 2: If the agent filed for permission in step 1, his project is investigated by an administrator who receives estimators $\hat{g}_A$ and $\hat{h}_A$ of the gains and losses associated with the project. The distribution of the estimators has a density $\phi_A(\hat{g}_A, \hat{h}_A; g, h, e_A)$ with $\frac{\partial \phi_A}{\partial g}$ being positive (negative) if and only if $\hat{g}_A$ is larger (smaller) than some $g_o(g)$ which is non-decreasing in g and the corresponding property for $\hat{h}_A$ and h. In words, the density $\phi_A(\hat{g}_A, \hat{h}_A; g, h)$ moves along the $\hat{g}_A$- and $\hat{h}_A$-axes as g and h, respectively, increase. Smaller

[6]This restriction ensures that the sets of g and h for which project owners file applications and appeal rejections, respectively for which third parties appeal approvals of projects are closed and contain all values above certain threshold values.

[7]Only if permission probabilities sharply decreased in g because larger g were correlated with much larger h, the gains g could become too *large* to deter the project owner from the application.

effort levels e_A result in a mean preserving spread of the distribution. As before, the administrator is benevolent with respect to the outcome of his decision (though not with respect to his effort e_A to come to this decision and with respect to being reversed by the courts). As the reader will see later, the administrator thus grants (denies) permission if his estimator of total harm accruing from the project is smaller (larger) than some critical value which is a strictly increasing function of his estimator of the gains accruing from the project, i.e. if $\hat{h}_A < (>) \tilde{h}_A(\hat{g}_A, e_A)$.

Step 3: If the administrator has denied permission in step 2, the applicant (owner of the project) decides on whether to lodge an appeal or not; if the administrator has granted permission, third parties make the corresponding decision. Lodging an appeal costs c_1 for applicants and c_3 for third parties. For simplicity, assume that the applicant has no information on the specific values of h and third parties have no information on the specific values of g. For similar reasons as argued for the project owner's decision to file for permission or not,[8] the applicant will lodge an appeal with the courts if and only if his individual gain is above some threshold level ($g > \bar{g}$) and third parties will lodge an appeal if and only if $\lambda h > \lambda \bar{h}$ where λ is the proportion of the harm which the most severely affected third party suffers.[9]

Step 4: If an appeal has been lodged by the owner of the project or third parties, a judge investigates the project and receives estimators $\hat{g}_J$ and $\hat{h}_J$ of the gains and losses associated with the project. The distribution of the estimators has a density $\phi_J(\hat{g}_J, \hat{h}_J; g, h, e_J)$ with the same properties as $\phi_A(\hat{g}_A, \hat{h}_A; g, h, e_A)$. In addition, the judge knows the decision of the administrator[10] and thus who lodged (and paid for) the appeal. As before, the judge is benevolent with respect to the outcome of his decision but not with respect to his effort e_J to come to this decision. As for the administrator, one will expect that the judge will grant permission if and only if the estimated harm is below some threshold level depending on the estimated gains from the project. However, the judge will condition his threshold level on the additional information he has, namely how the administrator decided

[8]... and with similar caveats ...

[9]The same caveat on fixed λ as expressed in footnote 2 is relevant again.

[10]In an extension of the model one could alternatively assume that the judge knows the administrator's estimators $\hat{g}_A$ and $\hat{h}_A$, for example from studying the reasons of the administrative decision. However, at least for Germany, the reasons of administrative decisions hardly give any clue on the evaluated gains and losses from the projects under decision (cf. Hofmann and von Wangenheim (2002)).

and who appealed the decision. Her condition to grant (deny) permission is thus $\hat{h}_J < (>) \tilde{h}^a_J(\hat{g}_J, e_J)$ if the administrator has approved the project, and $\hat{h}_J < (>) \tilde{h}^r_J(\hat{g}_J, e_J)$ if the administrator has rejected the project.

The gross payoff of the applicant is g if his project is eventually approved and 0 if not (be it due rejection of his application or due to not filing an application). From this gross payoff one has to deduct c_o if he files an application and, in addition, c_1 if he appeals a negative decision of the administrator. For the relevant third party[11] the gross payoff is $-\lambda h$ if the project is eventually approved and 0 if not. From this gross payoff one has to deduct c_3 if the third party appeals a positive decision of the administrator.

4.2 Informal Argument

As indicated in the description of the game played for each project, courts will grant permission to all projects for which the signal on the harm of the project is sufficiently low compared to the signal on its gains. What exactly "sufficiently low" means depends not only on the size of the gains but also on the courts' prior information on the project. This prior information stems from the decision of the administration and from the fact that the losing party has appealed this decision. Obviously, these two bits of information are going in opposite directions. The administrative decision indicates that the administrator has had some evidence (his estimates of gains and harm associated with the project) in the direction of the decision actually taken. The appeal of the losing party indicates that the party's information on the gains and harm associated with the project go in the opposite direction of the administrator's estimators.

Depending on how reliable[12] the competent judge thinks these two countervailing bits of information to be, she will tend to follow the administrative decision or tend to deviate from it. Of course, she will not completely rely on either bit of information, but allow for suprseding information from the estimators of gains and harm of the project derived in her own investigation of the project. If she did not, either the problem described by Shavell (cf. page 17 above) would reoccur and thus the prior information would be worthless, because the party losing in the administrative procedure would always

[11] i.e. the third party who suffers the largest part of the harm

[12] Here and in the following text including the formal analysis, I use the words "reliable" and "reliability" in an untechnical sense.

appeal independently of the party's information on the case. Or the judge would deprive herself of the possibility to use her own information even in cases were it is very reliable, because no administrative decision would ever be appealed.

Hence, the judge will use her prior information only to adjust the function of critical harm levels below which she will approve a project and above which she will reject it. If the information implied in the appeal of a party is more reliable than the information implied in the administration's decision, the judge will be reluctant to uphold the administrative decision unless her own estimators of harm and gains of the project strongly indicate that she should confirm. This implies that the judge may have estimators of the harm and gains for which he will reverse the administrative decision independently of its content if her own estimators are unclear as to the merits of the project. Of course, if the judge's own information is very reliable, the range of own estimators for which she will always overturn the administrative decision is very small.

In the reverse case, i.e. when the information implied in the administrator's decision is far more reliable than the information inferable from the fact of the appeal, the judge will tend to uphold the administrative decision. There now will be a range of gain and harm estimators of the judge for which she will uphold the administrative decision independently of its content. She will reverse the administrator's decision only if her own information very strongly goes in the other direction. This would be in accordance with the doctrine of judicial deference to administrative decisions, while the situation described before conflicts with this doctrine.

This may explain why courts seem to be not very consistent in their reliance on the doctrine. When the parties can easily detect the gains and harms of projects[13] but it is hard for the administration (and the courts) to find evidence on the social merits of the project,[14] then the fact that an appeal has been lodged may be strong information for the courts.[15] They will not grant any discretionary range to the administration and thus discard the doctrine of deference. Conversely, if administrators can easily observe gains

[13]More general: if they can easily detect the merits of their case.

[14]In other words: if the parties can easily observe gains and harm of the project but are unable to transmit this information to the adminsitration (and to the courts).

[15]One should note that the fact that a certain party has lodged the appeal is weak information if the appeals costs are small so that nearly every administrative decison is appealed.

and harms accruing from a project, the courts will take the administrative decsions as hard information and thus follow the doctrine of deference, i.e. grant the administration some range of discretion in which the courts own information will not be sufficient to overturn the administrative decisions.

This line of argument also concords with the dividing line between *unbestimmte Rechtsbegriffe* (unclear legal terms) and *Ermessensspielräume* (discretionary ranges) in German adminsitrative law. When the parties can more easily determine than the adminsitration whether their case is justified and when they can thus easily determine their chances in a court appeal, the law is interpreted as including unclear legal terms. When, on the other hand, the administration is as able as, or even more able than, the parties to estimate gains and harm of a project, German law is interpreted as granting discretionary ranges to the administration, courts follow a doctrine similar to judicial deference.[16]

Since the judges are assumed to be benevolent with respect to the content of their decisions, it is rather easy to argue that — apart from the resouces spent on the courts — appeals triggered judicial review is always welfare increasing, independently of the quality of judicial decision making. If the courts would never overturn admministrative decisions, this would be equivalent to granting infinite ranges of discretion to the administration. At the same time, this would ensure that social welfare would not be affected by the appeals triggered judicial review process. As the courts are assumed to be benevolent, they will deviate from this extreme form of deference if and only if this increases social welfare. In addition, one can easily see that this extreme form of deference is suboptimal from a social point of view (and that the courts will not follow it due to their benevolece). Even with very poor abilities of the courts, their information may sometimes be so extremely contrary to what the administration has decided that this information weighs stronger than the information included in the administrative decision.

4.3 Formal Argument

4.3.1 Equilibrium

The administrator's payoff U consists of an intrinsic and an extrinsic motivation referring to the content of his decisions and of his disutility of effort.

[16] Evidence from German law still has to be added.

In particular, his utility will increase by some fraction s_i of the welfare effects of his decisions (for simplicity, I assume that he does not take possible court reversals of his decisions into account insofar as he cares for the welfare effects of his decisions), will decrease by a constant s_x if one of his decisions is reversed, and will decrease by his disutility of effort u which depends on the total effort he exerts on making decisions on permissions. The administrator's payoff from deciding on a case is thus given by:

$$U = s_i(g - h) - s_x \, \rho - u(E_A) \tag{25}$$

where $\rho = 1$ if the courts reverse the administrator's decision and $\rho = 0$ if not, and E_A is the total effort exerted by the administrator.

The judge's payoff only consists of the intrinsic motivation and the disutility of effort for the reasons laid down in section 2 (page 14). Her payoffs are thus given by

$$V = s_i(g - h) - u(E_J) \tag{26}$$

where E_J is her total effort exerted on cases brought before her.

I assume that the choices of strategies, i.e. the choices of the sets of estimators for which administrators and judges approve or reject a project are simultaneous. This means in particular that administrators do not choose their strategies in order to induce a certain strategy choice of the courts nor *vice versa*.

One can then derive the Nash equilibrium of this game from the following argument. Note that this is not backward induction since the strategy choices are simultaneous. Nevertheless, it is most simple to start with the judge's decision on her effort per case and on her decision rules. Her expected payoff per case project for which the owner of the project may or may not file an application is given by:

$$
\begin{aligned}
\mathrm{E}V = & \\
& \left.\begin{aligned} s_i \iint_{\mathcal{G}\cap\mathcal{C}^a} \Bigg((g-h) \iint_{\mathcal{A}} \phi_A(\hat{g}_A, \hat{h}_A; g, h, e_A) \mathrm{d}\hat{h}_A \mathrm{d}\hat{g}_A \\ \cdot \iint_{\mathcal{J}^a} \phi_J(\hat{g}_J, \hat{h}_J; g, h, e_J) \mathrm{d}\hat{h}_J \mathrm{d}\hat{g}_J f(g,h) \Bigg) \mathrm{d}h\mathrm{d}g \end{aligned}\right\} \quad (27) \\
& \left.\begin{aligned} + s_i \iint_{\mathcal{G}\cap\mathcal{C}^r} \Bigg((g-h) \iint_{\overline{\mathcal{A}}} \phi_A(\hat{g}_A, \hat{h}_A; g, h, e_A) \mathrm{d}\hat{h}_A \mathrm{d}\hat{g}_A \\ \cdot \iint_{\mathcal{J}^r} \phi_J(\hat{g}_J, \hat{h}_J; g, h, e_J) \mathrm{d}\hat{h}_J \mathrm{d}\hat{g}_J f(g,h) \Bigg) \mathrm{d}h\mathrm{d}g \end{aligned}\right\} \quad (28) \\
& \left.\begin{aligned} -u \Bigg(e_J^a \iint_{\mathcal{G}\cap\mathcal{C}^a} \iint_{\mathcal{A}} \phi_J(\hat{g}_A, \hat{h}_A; g, h, e_A) \mathrm{d}\hat{h}_A \mathrm{d}\hat{g}_A f(g,h) \mathrm{d}h\mathrm{d}g \\ + e_J^r \iint_{\mathcal{G}\cap\mathcal{C}^r} \iint_{\overline{\mathcal{A}}} \phi_J(\hat{g}_A, \hat{h}_A; g, h, e_A) \mathrm{d}\hat{h}_A \mathrm{d}\hat{g}_A f(g,h) \mathrm{d}h\mathrm{d}g \Bigg) \end{aligned}\right\} \quad (29)
\end{aligned}
$$

where

- $\mathcal{A}$ and $\overline{\mathcal{A}}$ are the sets of all signals $\hat{g}_A$ and $\hat{h}_A$ for which the administrator will grant or deny, respectively, permission,
- $\mathcal{C}^a$ and $\mathcal{C}^r$ are the sets of all pairs of g and h for which third parties or the applicant, respectively, will appeal an approval or a rejection, respectively, of a project (will go to court, whence the "$\mathcal{C}$"),
- $\mathcal{J}^a$ and $\mathcal{J}^r$ are the sets of signals $\hat{g}_J$ and $\hat{h}_J$ for which the judge will grant permission given that the administration also granted permission and third parties appealed, or, respectively, for which the judge will grant permission given that the administration denied permission and the applicant appealed.

To understand the intuition behind the formal expression, consider lines (27) which describe the expected payoffs from intrinsic motivation for appeals in which the administrator approved the project. The outermost pair of integrals restricts the set of projects to those for which the applicant files an application and third parties appeal an administrative approval. Only these projects may come before the judge because of third party appeals and count for the balance of the welfare effects of her decisions. They come before her, if and only if the administrator approves them, which he does, if his signals are within the set $\mathcal{A}$. The probability of such a project coming before the judge is given for each (relevant) g–h–pair by the first inner double integral. The probability that a project not only comes before the judge by appeal of third parties but also counts for the balance of the welfare effects of her decisions, i.e. the probability that the judge approves such a project is given for each (relevant) g–h–pair by the second inner double interval. The product of these two probabilities is thus the probability that the judge approves a project from the set $\mathcal{G} \cap \mathcal{C}^a$. On gets the judge's expected intrinsic-motivation payoff from appeals in which the administrator approved the project by first weighting the welfare effects $g - h$ of all projects from the set $\mathcal{G} \cap \mathcal{C}^a$ by the latter probability and the density $f(g, h)$ with which someone ownes such a project, and then integrating over the set of all relevant projects $(\mathcal{G} \cap \mathcal{C}^a)$.

In a similar vein, lines (28) describe the expected payoffs from intrinsic motivation for appeals in which the administrator rejected the project. Finally, disutility of effort is a function of the effort exerted on cases in which a third party appealed and the probability that such an appeal is lodged per project plus the effort exerted on cases in which the owner of the project appealed and the propbability that such an appeal is lodged per project (line (29)).

Obviously, the judge will approve a project if and only if she estimates its gains to be large enough and its losses small enough. She will thus approve an administrative permissions if and only if her estimate $\hat{h}_J$ of harm associated with the project is smaller than a function $\tilde{h}^a_J(\hat{g}_J, e^a_J)$ of her estimate of the gains and her effort to improve her estimates. Similarly, she will approve a project which the administrator has rejected if and only if her estimate $\hat{h}_J$ of harm associated with the project is smaller than another function $\tilde{h}^r_J(\hat{g}_J, e^r_J)$ of her estimate of the gains and her effort to improve her estimates. This defines the limits of the inner intervals of the judge's expected payoffs so that equation (**??**) becomes:

$$
\begin{aligned}
\mathrm{E}V = \\
& s_i \iint\limits_{\mathcal{G}\cap\mathcal{C}^a} \Bigg((g-h) \iint\limits_{\mathcal{A}} \phi_A(\hat{g}_A, \hat{h}_A; g, h, e_A)\mathrm{d}\hat{h}_A \mathrm{d}\hat{g}_A \\
& \qquad \cdot \int_0^\infty \int_0^{\tilde{h}_J^a(\hat{g}_J, e_J^a)} \phi_J(\hat{g}_J, \hat{h}_J; g, h, e_J^a)\mathrm{d}\hat{h}_J \mathrm{d}\hat{g}_J f(g,h) \Bigg) \mathrm{d}h\mathrm{d}g \\
& + s_i \iint\limits_{\mathcal{G}\cap\mathcal{C}^r} \Bigg((g-h) \iint\limits_{\overline{\mathcal{A}}} \phi_A(\hat{g}_A, \hat{h}_A; g, h, e_A)\mathrm{d}\hat{h}_A \mathrm{d}\hat{g}_A \\
& \qquad \cdot \int_0^\infty \int_0^{\tilde{h}_J^r(\hat{g}_J, e_J^r)} \phi_J(\hat{g}_J, \hat{h}_J; g, h, e_J^r)\mathrm{d}\hat{h}_J \mathrm{d}\hat{g}_J f(g,h) \Bigg) \mathrm{d}h\mathrm{d}g \\
& - u \Bigg(e_J^a \iint\limits_{\mathcal{G}\cap\mathcal{C}^a} \iint\limits_{\mathcal{A}} \phi_J(\hat{g}_A, \hat{h}_A; g, h, e_A)\mathrm{d}\hat{h}_A \mathrm{d}\hat{g}_A f(g,h)\mathrm{d}h\mathrm{d}g \\
& \qquad + e_J^r \iint\limits_{\mathcal{G}\cap\mathcal{C}^r} \iint\limits_{\overline{\mathcal{A}}} \phi_J(\hat{g}_A, \hat{h}_A; g, h, e_A)\mathrm{d}\hat{h}_A \mathrm{d}\hat{g}_A f(g,h)\mathrm{d}h\mathrm{d}g \Bigg)
\end{aligned}
$$

The exact functions $\tilde{h}_J^a(\hat{g}_J, e_J^a)$ and $\tilde{h}_J^r(\hat{g}_J, e_J^r)$ are given by the threshold values for harm which maximize the expected welfare effect for every given gain estimate, hence by:

$$
0 \stackrel{!}{=} \frac{\partial \mathrm{E}V}{\partial \tilde{h}_J^a(\hat{g}_J, e_J^a)} = s_i \iint\limits_{\mathcal{G}\cap\mathcal{C}^a} (g-h) \iint\limits_{\mathcal{A}} \phi_A(\hat{g}_A, \hat{h}_A; g, h, e_A)\mathrm{d}\hat{h}_A \mathrm{d}\hat{g}_A \cdot \phi_J(\hat{g}_J, \tilde{h}_J^a(\hat{g}_J, e_J^a); g, h, e_J^a) f(g,h)\mathrm{d}h\mathrm{d}g \tag{30}
$$

and

$$
0 \stackrel{!}{=} \frac{\partial \mathrm{E}V}{\partial \tilde{h}_J^r(\hat{g}_J, e_J^r)} = s_i \iint\limits_{\mathcal{G}\cap\mathcal{C}^r} (g-h) \iint\limits_{\overline{\mathcal{A}}} \phi_A(\hat{g}_A, \hat{h}_A; g, h, e_A)\mathrm{d}\hat{h}_A \mathrm{d}\hat{g}_A \cdot \phi_J(\hat{g}_J, \tilde{h}_J^r(\hat{g}_J, e_J^r); g, h, e_J^r) f(g,h)\mathrm{d}h\mathrm{d}g \tag{31}
$$

From the assumptions on $\phi_J(\cdot)$ it immediately follows that $\tilde{h}^a_J(\hat{g}_J, e^a_J)$ and $\tilde{h}^r_J(\hat{g}_J, e^r_J)$ are unique functions for any given e^a_J and e^r_J and increase in their first arguments.

Optimal effort levels follow from equating the derivatives of EV with respect to e^a_J and e^r_J to zero. Their uniqueness depends on how exactly effort influences the distribution of the judge's estimators.

For given functions $\tilde{h}^a_J(\hat{g}_J, e^a_J)$ and $\tilde{h}^r_J(\hat{g}_J, e^r_J)$ of critical harm-gains-relationships, the losing party of the administrative decision process decides on lodging an appeal. Specifically, the expected payoff from lodging an appeal for the owner of a project is given by:

$$\begin{aligned} \mathrm{E}\Pi^r(g) = & -c_1 \\ & + g\frac{\int\limits_0^\infty \left[\int\limits_0^\infty \int\limits_0^{\tilde{h}^r_J(\hat{g}_J, e^r_J)} \phi_J(\hat{g}_J, \hat{h}_J; g, h, e^r_J)\mathrm{d}\hat{h}_J \mathrm{d}\hat{g}_J \iint\limits_{\overline{\mathcal{A}}} \phi_A(\hat{g}_A, \hat{h}_A; g, h, e_A)\mathrm{d}\hat{h}_A \mathrm{d}\hat{g}_A f(g,h) \right] \mathrm{d}h}{\int\limits_0^\infty \left[\iint\limits_{\overline{\mathcal{A}}} \phi_A(\hat{g}_A, \hat{h}_A; g, h, e_A)\mathrm{d}\hat{h}_A \mathrm{d}\hat{g}_A f(g,h) \right] \mathrm{d}h} \end{aligned} \tag{32}$$

where

$$\frac{\iint\limits_{\overline{\mathcal{A}}} \phi_A(\hat{g}_A, \hat{h}_A; g, h, e_A)\mathrm{d}\hat{h}_A \mathrm{d}\hat{g}_A f(g,h)}{\int\limits_0^\infty \left[\iint\limits_{\overline{\mathcal{A}}} \phi_A(\hat{g}_A, \hat{h}_A; g, h, e_A)\mathrm{d}\hat{h}_A \mathrm{d}\hat{g}_A f(g,h) \right] \mathrm{d}h}$$

is the marginal distribution of h for a given g and under the condition that the administrator has rejected the project.

As I assumed that an increase in g is not highly correlated with a sharp increase in h,[17] E$\Pi^r(g)$ strictly increases in g. Hence, the set $\mathcal{C}^r$ of g–h combinations for which the owner of a project sues against the administrative rejection of his plans is given by all g–h combinations for which g is above some threshold level $\bar{g}$ (independently of h). Note that the appellant includes the information embodied in the administration's rejection of his project in his estimation of his payoffs from appealing: he uses the aforementioned marginal conditional density to determine the now relevant distribution of h.

Likewise, the third party which suffers most of the harm of a project with

[17] Cf. the assumptions on $f(g,h)$ in step 0 of the description of the game.

total harm of any given h, receives an expected payoff of

$$\mathrm{E}\Pi^a(h) = -c_3 + \lambda h \frac{\int\limits_{\mathcal{G}} \left[\int\limits_0^\infty \int\limits_{\tilde{h}_J^a(\hat{g}_J, e_J^a)}^\infty \phi_J(\hat{g}_J, \hat{h}_J; g, h, e_J^a) \mathrm{d}\hat{h}_J \mathrm{d}\hat{g}_J \iint\limits_{\mathcal{A}} \phi_A(\hat{g}_A, \hat{h}_A; g, h, e_A) \mathrm{d}\hat{h}_A \mathrm{d}\hat{g}_A f(g,h) \right] \mathrm{d}g}{\int\limits_{\mathcal{G}} \left[\iint\limits_{\mathcal{A}} \phi_A(\hat{g}_A, \hat{h}_A; g, h, e_A) \mathrm{d}\hat{h}_A \mathrm{d}\hat{g}_A f(g,h) \right] \mathrm{d}g}$$

from lodging an appeal after the administrator has approved the project. Note that

$$\frac{\iint\limits_{\mathcal{A}} \phi_A(\hat{g}_A, \hat{h}_A; g, h, e_A) \mathrm{d}\hat{h}_A \mathrm{d}\hat{g}_A f(g,h) \mathrm{d}g}{\int\limits_{\mathcal{G}} \left[\iint\limits_{\mathcal{A}} \phi_A(\hat{g}_A, \hat{h}_A; g, h, e_A) \mathrm{d}\hat{h}_A \mathrm{d}\hat{g}_A f(g,h) \right] \mathrm{d}g}$$

is the marginal distribution of g for a given h and under the condition that the administrator has approved the project. As $\mathrm{E}\Pi^a(h)$ is also a strictly increasing function,[18] the set $\mathcal{C}^a$ consists of all g–h combinations for which h is above some threshold level $\bar{h}$ (independently of g). Again it is noteworthy that the appellant relies on the relevant marginal conditional density of h, and thus makes use of the information that the administrator has approved the project.

One can now turn to the administrator's behavior. His expected utility per project for which the owner may file an application or not, is given by:

$$\begin{aligned}
\mathrm{E}U = \; & s_i \iint\limits_{\mathcal{G}} (g-h) \iint\limits_{\mathcal{A}} \phi_A(\hat{g}_A, \hat{h}_A; g, h, e_A) \mathrm{d}\hat{h}_A \mathrm{d}\hat{g}_A f(g,h) \mathrm{d}h \mathrm{d}g \\
& - s_x \iint\limits_{\mathcal{G}\cap\mathcal{C}^r} \left[\iint\limits_{\overline{\mathcal{A}}} \phi_A(\hat{g}_A, \hat{h}_A; g, h, e_A) \mathrm{d}\hat{h}_A \mathrm{d}\hat{g}_A \int\limits_0^\infty \int\limits_0^{\tilde{h}_J^r(\hat{g}_J, e_J)} \phi_J(\hat{g}_J, \hat{h}_J; g, h, e_J) \mathrm{d}\hat{h}_J \mathrm{d}\hat{g}_J f(g,h) \right] \mathrm{d}h \mathrm{d}g \\
& - s_x \iint\limits_{\mathcal{G}\cap\mathcal{C}^a} \left[\iint\limits_{\mathcal{A}} \phi_A(\hat{g}_A, \hat{h}_A; g, h, e_A) \mathrm{d}\hat{h}_A \mathrm{d}\hat{g}_A \int\limits_0^\infty \int\limits_{\tilde{h}_J^a(\hat{g}_J, e_J)}^\infty \phi_J(\hat{g}_J, \hat{h}_J; g, h, e_J) \mathrm{d}\hat{h}_J \mathrm{d}\hat{g}_J f(g,h) \right] \mathrm{d}h \mathrm{d}g \\
& - u \left(e_A \iint\limits_{\mathcal{G}} f(g,h) \mathrm{d}h \mathrm{d}g \right)
\end{aligned}$$

[18]... unless an increase in h is highly correlated with a sharp increase in g,

The first line of the expression describes the administrator's intrinsic motivation, the second and third lines describe the administrator's expected disutility from being reversed after having rejected and, respectively, approved the project, and the last line describes his disutility of effort.

As for the judge, it is obvious that the administrator will approve a project only if he estimates its gains to be large enough and its losses small enough. He will thus approve a project if and only if his estimate $\hat{h}_A$ of harm associated with the project is smaller than a function $\tilde{h}_A(\hat{g}_A, e_A)$ of his estimate of the gains and his effort to improve her estimates. This defines the limits of the inner intervals of the administrator's expected payoffs so that equation (33) becomes:

$$
\begin{aligned}
\mathrm{E}U = \quad & s_i \iint_{\mathcal{G}} (g-h) \int_0^{\infty} \int_0^{\tilde{h}_A(\hat{g}_A, e_A)} \phi_A(\hat{g}_A, \hat{h}_A; g, h, e_A) \mathrm{d}\hat{h}_A \mathrm{d}\hat{g}_A f(g,h) \mathrm{d}h \mathrm{d}g \\
& - s_x \iint_{\mathcal{C}^r} \left[\int_0^{\infty} \int_{\tilde{h}_A(\hat{g}_A, e_A)}^{\infty} \phi_A(\hat{g}_A, \hat{h}_A; g, h, e_A) \mathrm{d}\hat{h}_A \mathrm{d}\hat{g}_A \int_0^{\infty} \int_0^{\tilde{h}^r_J(\hat{g}_J, e_J)} \phi_J(\hat{g}_J, \hat{h}_J; g, h, e_J) \mathrm{d}\hat{h}_J \mathrm{d}\hat{g}_J f(g,h) \right] \mathrm{d}h \mathrm{d}g \\
& - s_x \iint_{\mathcal{C}^a} \left[\int_0^{\infty} \int_0^{\tilde{h}_A(\hat{g}_A, e_A)} \phi_A(\hat{g}_A, \hat{h}_A; g, h, e_A) \mathrm{d}\hat{h}_A \mathrm{d}\hat{g}_A \int_0^{\infty} \int_{\tilde{h}^a_J(\hat{g}_J, e_J)}^{\infty} \phi_J(\hat{g}_J, \hat{h}_J; g, h, e_J) \mathrm{d}\hat{h}_J \mathrm{d}\hat{g}_J f(g,h) \right] \mathrm{d}h \mathrm{d}g \\
& - u \left(e_A \iint_{\mathcal{G}} f(g,h) \mathrm{d}h \mathrm{d}g \right) \qquad (33)
\end{aligned}
$$

Finally, consider the payoff of the owner of a project when making his decision whether to file for an application or not. If his gains from the project is any given g and he files, his expected payoff is:

$$
\begin{aligned}
\mathrm{E}\Pi^o(g) = \frac{g}{\int\limits_0^\infty f(g,h)\mathrm{d}h} \Bigg(& \int\limits_0^{\bar{h}}\int\limits_0^\infty\int\limits_0^{\tilde{h}_A(\hat{g}_A,e_A)} \phi_A(\hat{g}_A,\hat{h}_A;g,h,e_A)\mathrm{d}\hat{h}_A\mathrm{d}\hat{g}_A f(g,h)\mathrm{d}h + \\
& \int\limits_{\bar{h}}^\infty\int\limits_0^\infty\int\limits_0^{\tilde{h}_A(\hat{g}_A,e_A)} \phi_A(\hat{g}_A,\hat{h}_A;g,h,e_A)\mathrm{d}\hat{h}_A\mathrm{d}\hat{g}_A \int\limits_0^\infty\int\limits_0^{\tilde{h}_J(\hat{g}_J,e_J)} \phi_J(\hat{g}_J,\hat{h}_J;g,h,e_J)\mathrm{d}\hat{h}_J\mathrm{d}\hat{g}_J f(g,h)\mathrm{d}h \Bigg) \\
& + \frac{\max\big(\mathrm{E}\Pi^r(g),0\big)}{\int\limits_0^\infty f(g,h)\mathrm{d}h} \int\limits_0^\infty\int\limits_0^\infty\int\limits_{\tilde{h}_A(\hat{g}_A,e_A)}^\infty \phi_A(\hat{g}_A,\hat{h}_A;g,h,e_A)\mathrm{d}\hat{h}_A\mathrm{d}\hat{g}_A f(g,h)\mathrm{d}h \qquad (34)
\end{aligned}
$$

The first two lines of expression (34) describe the expected payoff from being granted permission without any appeal or after a third party appeal, respectively, and the third line describes the expected payoff from being granted permission only after an own appeal. Note that $f(g,h)\Big/\int\limits_0^\infty f(g,h)\mathrm{d}h$ is the marginal distribution of h for any given g.

Again it is easy to see that $\mathrm{E}\Pi^o(g)$ is an increasing function for not too extreme density functions $f(g,h)$, so that the set of all projects for which an application is filed is given by all combinations of g and h for which g is above some threshold level $\underline{g}$.

4.3.2 Judicial Deference?

As argued in the previous section, courts will usually not choose the same functions $\tilde{h}^a_J(\hat{g}_J,e^a_J)$ and $\tilde{h}^r_J(\hat{g}_J,e^r_J)$ for appeals of third parties and applicants. It depends on the qualities of the information provided by the administrative decision and the fact that an appeal has been lodged whether $\tilde{h}^a_J(\hat{g}_J,e^a_J) > \tilde{h}^r_J(\hat{g}_J,e^r_J)$ or $\tilde{h}^a_J(\hat{g}_J,e^a_J) < \tilde{h}^r_J(\hat{g}_J,e^r_J)$, i.e. whether the courts leave the administrative decision unaffected if in doubt or reverse the administrative decision if in doubt.

To understand more precisely what is meant by the qualities of the information provided by the administratove decision and the fact that an appeal has been lodged reconsider equations (30) and (31). These definitions of the functions of critical gain-harm-relationships for the two kinds of appeals differ in two respects: the set of g–h–combinations for which an appeal is filed ($\mathcal{C}^a$ versus $\mathcal{C}^r$) and the probabilities with which projects of any given gains and harms are approved or rejected ($\iint\limits_{\mathcal{A}} \phi_A(\hat{g}_A, \hat{h}_A; g, h, e_A)\mathrm{d}\hat{h}_A \mathrm{d}\hat{g}_A$ versus $\iint\limits_{\overline{\mathcal{A}}} \phi_A(\hat{g}_A, \hat{h}_A; g, h, e_A)\mathrm{d}\hat{h}_A \mathrm{d}\hat{g}_A$).

To see how various specifications of these two bits of information influence the judge's decision rule, consider two extreme cases with the simplification that the judge's effort is the same for both types of appeals ($e^r_J = e^a_J$).

Case 1: Suppose that the administrator's signals provide hardly any information on g and h, i.e. that the joint distribution of his estimators are (nearly) independent of the true values of g and h. Further suppose that the costs of lodging an appeal are considerable so that $\mathcal{C}^a$ excludes all projects with small harm h while $\mathcal{C}^r$ excludes all projects with small gains. Finally suppose that the judge's signals are not too reliable. Then the optimality condition (30) can only be satisfied, if for any given signal $\hat{g}_J$ on the gains of the project, the judge grants permission only to projects with low signals $\hat{h}_J$ for the damages: otherwise positive values of $g - h$ (which can only result from relatively large g, since small h are not within $\mathcal{C}^a$) would not be given sufficient weight in the outer integral in condition (30) to keep the value of the integral zero. Hence, $\tilde{h}^a_J(\hat{g}, e^a_J)$ must be small for every given $\hat{g}_J$.

Now suppose that $\tilde{h}^r_J(\hat{g}, e^r_J)$ were the same as $\tilde{h}^a_J(\hat{g}, e^a_J)$. Then there is the set $\mathcal{G} \cap \mathcal{C}^a \cap \mathcal{C}^r$ which is part of the support of the outer integrals in both conditions (30) and (31) and for which the arguments of these integrals are the same. However, the rest of the support of the outer integral in equation (30) (i.e. $\mathcal{G} \cap \mathcal{C}^a \backslash \mathcal{C}^r$) comprises mainly pairs of g and h for which $g < h$ and the rest of the support of the outer integral in equation (31) (i.e. $\mathcal{G} \cap \mathcal{C}^r \backslash \mathcal{C}^a$) comprises mainly pairs of g and h for which $g > h$. Thus, unless the judge's signals are very precise, the value of the rest of the support of the outer integral in equation (30) will tend to be negative and the rest of the support of the outer integral in equation (31) will tend to be positive; at least the former will be smaller than the latter. Thus, the value of zero of the integral in (30) implies a positive value of the intergral in equation (31). To reduce the value of the latter integral to zero, the judge must increase $\tilde{h}^r_J(\hat{g}, e^r_J)$ above the value of $\tilde{h}^a_J(\hat{g}, e^a_J)$ for every $\hat{g}_J$, i.e. she must approve projects for which she received signals indicating a lower, possibly negative welfare effect.

The result $\tilde{h}^r_J(\hat{g}, e^r_J) > \tilde{h}^a_J(\hat{g}, e^a_J)$ implies that the judge approves projects with larger estimators of harm for any given estimate of gains if the administrator had rejected the project than if he had approved it. The same estimates of harm and gains may thus induce the judge to approve the project of the administrator had rejected it and to reject the project if the judge had approved the it. For some estimates of gains and harm the judge thus reverses the administrator's decision independently of how the administrator had decided. In this sense, one could hence say for some estimators of the judge that the administrator can only have done wrong.

Case 2: Suppose that the administrator's signals are very reliable, i.e. provide a lot of information on g and h. The probability mass of the joint distribution of his estimators are very much concentrated around the true values of g and h. Further suppose that the costs of lodging an appeal are small so that $\mathcal{C}^a$ includes projects even with small harm h (only the smallest values of h are excluded) while $\mathcal{C}^r$ includes projects even with small gains (again only the smallest values of g are excluded). $\mathcal{C}^a$ and $\mathcal{C}^r$ thus hardly differ — the information on who has lodged the appeal is of hardly any use for the judge. Finally, again suppose that the judge's estimators are not too reliable.

The crucial difference in the expressions in equations (30) and (31) — apart from the ϕ_J–functions — then are the inner integrals. The value of the inner integral in equation (30) strongly increases in $g-h$, while the value of the inner integral in equation (31) strongly decreases in $g-h$. Hence larger welfare gains $g-h$ tend to have high weights in the integral in equation (30) and low weights in the integral in equation (31). The reverse is true for smaller welfare gains. Hence, if $\tilde{h}^r_J(\hat{g}, e^r_J)$ were equal to $\tilde{h}^a_J(\hat{g}, e^a_J)$, for any given $\hat{g}_J$, then the value of the integral in equation (30) would be larger than value of the integral in equation (31). Since both have to be zero in the judge's optimal behavior, $\tilde{h}^r_J(\hat{g}, e^r_J)$ has to be smaller than $\tilde{h}^a_J(\hat{g}, e^a_J)$ since then $\phi_J(\hat{g}_J, \tilde{h}^r_J, g, h, e^r_J)$ is smaller than $\phi_J(\hat{g}_J, \tilde{h}^a_J, g, h, e^a_J)$ for larger h (and thus smaller $g-h$) and smaller for smaller h and thus larger $g-h$. This will eventually offset the differences in size of the inner integral in equations (30) and (31).

The result $\tilde{h}^r_J(\hat{g}, e^r_J) < \tilde{h}^a_J(\hat{g}, e^a_J)$ is opposit to the result for case 1. The result of case 2 implies that if the administrator has approved a project, the judge deems estimators of harm sufficiently low to approve the project which would be too high had the administrator rejected the project. For some estimates of gains and harm the judge thus approves the administrator's decision independently of how the administrator had decided. The judge thus leaves some range of discretion to the adminsitrator: only if the judge's estimators indicate that the administrator's decision was crudely wrong, she will overturn the administrative decision.

One could construct further cases, but the additional insights would be minor or directly intuitive. If the information provided by the appeal by one particular party and by the administrative decision are equally reliable, they will offset each other in equations (30) and (31) so that the optimal function $\tilde{h}^r_J(\hat{g}, e^r_J)$ will be equal to the optimal function $\tilde{h}^a_J(\hat{g}, e^a_J)$. If the reliability of both is small compared to the reliability of the judge's own estimators, the above argument will tend to continue to hold true, but the difference between $\tilde{h}^r_J(\hat{g}, e^r_J)$ and $\tilde{h}^a_J(\hat{g}, e^a_J)$ will become smaller.

In the model developed in this section, the courts will thus follow the doctrine of judicial deference if and only if the information which the courts can infer from the administrative decision is relatively more reliable than the information they can infer from the fact that (and which) party appealed the administrative decision. I have already elaborated on the consequences of this result for related legal questions in the informal discussion of the model (page 22).

4.3.3 Welfare Effects of Judicial Review

Having discussed the behavior of judges in the previous two subsections, the question remains whether judicial review, be it appeals triggered or not, increases social welfare or not. The answer to this question does not follow directly from the results in section 2.2, since contrary to there, appeals of correct administrative decisions now may occur as well. The expected effect of the entire decision procedure on social welfare of each case in which an owner of a project may file for permission is given in extension of the two-projects case by

$$W = \int_{\underline{g}}^{\infty}\int_{-\infty}^{\infty}\big((g-h)p(g,h) - c_A - p_{\text{appeal}}(g,h)c_J\big)\mathrm{d}h\mathrm{d}g \qquad (35)$$

where $p(g,h)$ is the probability that the project passes the combination of the administrative and the judicial decision procedure, $p_{\text{appeal}}(g,h)$ is the probability of an appeal, c_J are the resource costs of an appeal to society and c_A are the corresponding costs of the administrative decision making.

To see how the appeals process changes changes social welfare, consider first the cases in which project approving decisions of the administration are appealed. Leaving the resource costs of judicial review aside, the welfare effect of judicial review in these cases may be written as:

$$\Delta W^a = \iint_{\mathcal{G}\cap\mathcal{C}^a}(g-h)\iint_{\mathcal{A}}\phi_A(\hat{g}_A,\hat{h}_A;g,h,e_A)\mathrm{d}\hat{h}_A\mathrm{d}\hat{g}_A \cdot \left(\int_0^\infty\int_0^{\tilde{h}_J^a(\hat{g}_J,e_J^a)}\phi_J(\hat{g}_J,\hat{h}_J;g,h,e_J^a)\mathrm{d}\hat{h}_J\mathrm{d}\hat{g}_J - 1\right) f(g,h)\mathrm{d}h\mathrm{d}g \tag{36}$$

The limits of the outer integral restrict the discussion to the projects for which an application is filed and which are appealed in case of administrative approval. The first factor within the outer integral gives the social values of these projects. The first inner integral is the probability for each of these projects that the administration approves it. The last factor ($f(g,h)$) is the frequency with which applications for such projects may be filed. What is crucial is the term in parentheses. If the courts decide on such such projects, the probability that they are approved and thus have an influence on social welfare is given by the integral within the parentheses. If no appeals process existed, the administrative decision on them would remain unaffected: they would be approved with certainty (the probability that the competent administrator approves is already represented by the first inner interval). The term in parentheses thus describes how the probability of approval of such projects changes after and given the administrative decision, if the appeals process becomes possible. Taking the first derivative with respect to $\tilde{h}_J^a(\hat{g}_J,e_J^a)$ yields the same expression as in equation (30), except for the proportionality factor s_i. This is only natural, since judges are benevolent with respect to the content of their decisions. Due to her benevolence, the judge chooses $\tilde{h}_J^a(\hat{g}_J,e_J^a)$ in the socially optimal way.

The question remains whether this socially optimal $\tilde{h}_J^a(\hat{g}_J,e_J^a)$ induces a welfare gain as compared to no judicial review at all. To see that the answer must be afirmative, suppose for the moment that $\tilde{h}_J^a(\hat{g}_J,e_J^a)$ were very large (tending towards infinity) for every given $\hat{g}$. Since this would imply judicial approval of nearly all projects which had been approved by the administration, the term in parentheses in equation (36) then increases toward zero. This is reinforced by the fact that $\mathcal{C}^a$ becomes ever smaller: only the cases with the very large harms will be appealed since the others have hardly any chance of being rejected by the courts. This also implies that only negative values of $g-h$ get relevant weights in the outer integrals in equations (30) and (36). Since the equilibrium function $\tilde{h}^a_J(\hat{g}_J, e^a_J)$ must be finite (mind the aforementioned SHAVELLian problem) and is welfare maximizing, the welfare effect of the equilibrium $\tilde{h}^a_J(\hat{g}_J, e^a_J)$ must thus be larger than zero.

By a parallel argument, one can easily show that the welfare effect of judicial review in cases in which project rejecting decisions of the administration are appealed must also be larger than zero. Hence, apart from the resource costs of judicial review, apeals triggered court reviews increase social welfare. This result extends the insight of section 2.2 to the case of a continuous set of possible projects.

The incentives effect of stochastic review extends naturally from the case of only two types of applications discussed in section 2.3 to the case of a continous set of possible projects: If the term s_x which represents the relevance of extrinsic motivation may be increased in a costless way, then it will always be possible to turn the welfare effect of judicial control positive, even if the courts' estimates of gains and harm are only weekly correlated with the true gains and harm.

Given the extensive formal apparatus needed to prove superiority of expert control for the case of only two types of projects (section 3), the question whether poor control is better than good control in the case of a continous set of possible projects must remain for for future research.

5 Conclusions

In this paper, I have shown that judicial control of administrative decisions may be welfare improving even if courts are substantially more likely to approve welfare reducing projects and reject welfare enhancing projects than administrators. In this sense, non-expert judicial control of expert administrative decisions may be worthwhile. However, at least for the simple model of only two alternative project types in the SHAVELLian tradition, I could also show that expert control would even be better than non-expert control. In the framework of the extension of the simple model to the case of a two-dimensional continuum of projects, I discussed how the courts' willingness to approve projects depends on the prior administrative decisions. It turned out that courts which are benevolent with respect to the content of their decisions follow the doctrine of judicial deference if and only if it is not easier for the parties to observe the true welfare effects of the projects under discussion. If parties can easily observe the welfare qualities of projects but the administration cannot, then courts tend to be hypercritical with respect to the administrative decisions. For some little conclusive information sets of the courts, they will reject the administrative decision independently of its content and replace it with its contrary.

The approach still leaves open a bundle of variations of the model which may be worthwhile to investigate more closely. In section 4, I have assumed that the parties know nothing about the gains or harms a project inflicts on the other party in the specific case. One could alternatively assume that the parties, before deciding on filing an application or lodging an appeal, derive estimators in the same way as administrators and judges do. While the analysis would become more complicated, I conjecture that the results of that section would remain the same. Only the strict borderlines of the sets of projects for which an application is filed or an appeal lodged would fade. Due to the stochasticity in the relevant estimators, every project could be within the set of applications and appeals of both types, though the true values of gains and harm would still determine the probabilities of being in these sets.

One could go even one step further and assume that the parties perfectly know both gains and harms of each project. Still the information provided by the appeals (or applications) would not be perfectly reliable because some applications for projects with negative welfare effects have to be filed, some administrative rejections of projects with (small) negative welfare effects have to be appealed and some administrative approvals of projects with (small) positive welfare effects have to be appealed in equilibrium. Otherwise, one would run into the SHAVELLian problem of non-existing Nash equilibria in pure strategies.

In another extension of the model, one could assume that courts are not only informed about the binary value of the administrative decision, but also on their findings of gains and harm of a project. Again, the central arguments would not change: whether the courts grant discretionary ranges to the administration or repeal administrative decisions in case of unclear court estimators still depends on the relative quality of the information the courts can infer from the administrative decision and the fact that one party appealed that decision. Only the set of judicial estimators which cast sufficient doubt on their reliability may vary with the absloute size of the difference between the administrator's estimates of gains and harm.

Variations concerning the benevolence of administrators and judges with respect to the content of the decisions might change the results substantially. If lacking expertise not only results in higher error probabilities but also in biased error probabilities, non-expert control may reinforce the problem of imperfect administrators at least insofar as the argument rests on the incentives effect.

Last but not least, the extension already mentioned at the end of section 4, investigation of variations in relative expertise of administrators and judges suggests itself.

A Exact values of numerators of equations (20) and (21)

The exact values of numerators of equations (20) and (21) are given by:

$$\begin{vmatrix} \frac{\partial^2 \mathrm{E}U_n}{\partial e_A \partial M_A} & \frac{\partial^2 \mathrm{E}U_n}{\partial e_A \partial e_J} \\ \frac{\partial^2 \mathrm{EV}_n}{\partial e_J \partial M_A} & \frac{\partial^2 \mathrm{EV}_n}{\partial e_J^2} \end{vmatrix} = -\frac{e_A\, q_A(e_A)\, n^3}{M_A^3 M_J}\, u''\Big(e_A \frac{n}{M_A}\Big) \cdot \left(s_i \tilde{w}\, q_J''(e_J) + u''\Big(e_J \frac{n\, q_A(e_A)}{M_J}\Big) \frac{n\, q_A(e_A)}{M_J}\right) < 0$$

$$\begin{vmatrix} \frac{\partial^2 \mathrm{E}U_n}{\partial e_A \partial M_J} & \frac{\partial^2 \mathrm{E}U_n}{\partial e_A \partial e_J} \\ \frac{\partial^2 \mathrm{EV}_n}{\partial e_J \partial M_J} & \frac{\partial^2 \mathrm{EV}_n}{\partial e_J^2} \end{vmatrix} = -\frac{s_x\, e_J\, q_A(e_A)^2\, q_A'(e_A)\, q_J'(e_J)\, n^3}{M_J^2}\, u''\Big(e_J \frac{n\, q_A(e_A)}{M_A M_J^3}\Big) < 0$$

$$\begin{vmatrix} \frac{\partial^2 \mathrm{E}U_n}{\partial e_A^2} & \frac{\partial^2 \mathrm{E}U_n}{\partial e_A \partial M_A} \\ \frac{\partial^2 \mathrm{EV}_n}{\partial e_A \partial e_J} & \frac{\partial^2 \mathrm{EV}_n}{\partial e_J \partial M_A} \end{vmatrix} = \frac{e_A\, e_J\, q_A(e_A)\, q_A'(e_A)\, n^4}{M_A^3 M_J^2}\, u''\Big(e_A \frac{n}{M_A}\Big) u''\Big(e_J \frac{n\, q_A(e_A)}{M_J}\Big) < 0$$

$$\begin{vmatrix} \frac{\partial^2 \mathrm{E}U_n}{\partial e_A^2} & \frac{\partial^2 \mathrm{E}U_n}{\partial e_A \partial M_J} \\ \frac{\partial^2 \mathrm{EV}_n}{\partial e_A \partial e_J} & \frac{\partial^2 \mathrm{EV}_n}{\partial e_J \partial M_J} \end{vmatrix} = -\frac{e_J\, q_A(e_A)^2\, n^3}{M_J^3 M_A} u''\Big(e_J \frac{n\, q_A(e_A)}{M_J}\Big) \cdot \left(\big(s_i \tilde{w} + s_x (1 - q_J(e_J))\big) q_A''(e_A) + u''\Big(e_A \frac{n}{M_A}\Big) \frac{n}{M_A}\right) < 0$$

and thus have negative signs as claimed in the text.

References

Bawn, Kathleen (1997), 'Choosing Strategies to Controle the Bureaucracy: Statutory Constraints, Oversight, and the Committee System', *13 Journal of Law, Economics, and Organization*, 101–126

Hofmann, Ekkehard and von Wangenheim, Georg (2002), 'Trade-Secrets versus Cost Benefit Analysis' *22 International Review of Law and Economics* forthcoming

McCubbins, Mathew D. and Schwartz, T. (1984), 'Congressional Oversight Overlooked: Policy Patrols vs. Fire Alarms', *28 American Journal of Political Science*, 165–179

Samuelson, Larry (1997), *'Evolutionary Games and Equilibrium Selection'* Cambridge, Mass, et al.: MIT Press

Shavell, Steven (1995), 'The Appeals Process as a Means of Error Correction' *24 Journal of Legal Studies*, 379–426

von Wangenheim, Georg (2000), 'Production of Legal Rules by Agencies and Bureaucracies', in: Boudewijn Bouckaert and Gerrit De Geest (eds): *Encyclopedia of Law and Economics, Vol. V*, Aldershot: Edward Elgar, 559–586

Weingast, Barry R. (1984), 'The Congressional-Bureaucratic System: A Principal Agent Perspective (With Applications to the SEC)', *44 Public Choice*, 147-191

Contents