Interdisciplinary Public Policy Mainz

IPP

Gutenberg School of Management and Economics

& Research Unit "Interdisciplinary Public Policy"

Discussion Paper Series

What Explains 'Generosity' in the Public Financing of High-Tech Drugs?

An Empirical Investigation for 25 OECD Countries and 11 Controversial Drugs

Katharina Böhm, Claudia Landwehr and Nils Steiner

Discussion paper number 1708

Johannes Gutenberg University Mainz
Gutenberg School of Management and Economics
Jakob-Welder-Weg 9
55128 Mainz
Germany
wiwi.uni-mainz.de

Contact details:

Katharina Böhm
Department of Political Science
Johannes Gutenberg University Mainz
Jakob-Welder-Weg 12
55128 Mainz

boehm@politik.uni-mainz.de

Claudia Landwehr
Department of Political Science
Johannes Gutenberg University Mainz
Jakob-Welder-Weg 12
55128 Mainz

landwehr@politik.uni-mainz.de

Nils D. Steiner
Institut für Politikwissenschaft
Johannes Gutenberg University Mainz
Jakob-Welder-Weg 12
55128 Mainz

steiner@politik.uni-mainz.de

What Explains 'Generosity' in the Public Financing of High-Tech Drugs?

An Empirical Investigation for 25 OECD Countries and 11 Controversial Drugs[1]

Katharina Böhm (boehm@politik.uni-mainz.de)

Claudia Landwehr (landwehr@politik.uni-mainz.de)

Nils D. Steiner (steiner@politik.uni-mainz.de)

Abstract:

In times of increasing cost pressures public health care systems in the OECD countries face the question whether and to which extent new high-tech drugs are to be financed within their public health care systems. Systematic empirical research that tries to explain across-country variation in these coverage decisions is, however, almost non-existent. We analyze an original dataset that contains coverage decisions for 11 controversial drugs in 25 OECD countries via multilevel modeling. Our results indicate that regulations to what extent a pharmaceutical is publicly financed are unrelated to wealth and general expenditure levels for health care, while societal health care systems tend to be more generous. By taking into account that rationing decisions have been (at least partially) delegated to specialized agencies in all of the countries under investigation, we also uncover suggestive evidence that institutional characteristics of the underlying decision processes matter systematically for coverage decisions.

[1] This is a postprint version of the following article: Böhm, Katharina / Landwehr, Claudia / Steiner, Nils 2014: What explains generosity in the public financing of high-tech drugs? An empirical investigation of 25 OECD countries and 11 controversial drugs (mit Katharina Böhm und Nils Steiner), *Journal Of European Social Policy*, 24 (1), 39-55.

Keywords:
Delegation; independent regulatory agencies; health care; priority setting; multilevel analysis

Acknowledgements:
The authors thank Kai Arzheimer for helpful comments on an earlier version of this paper.

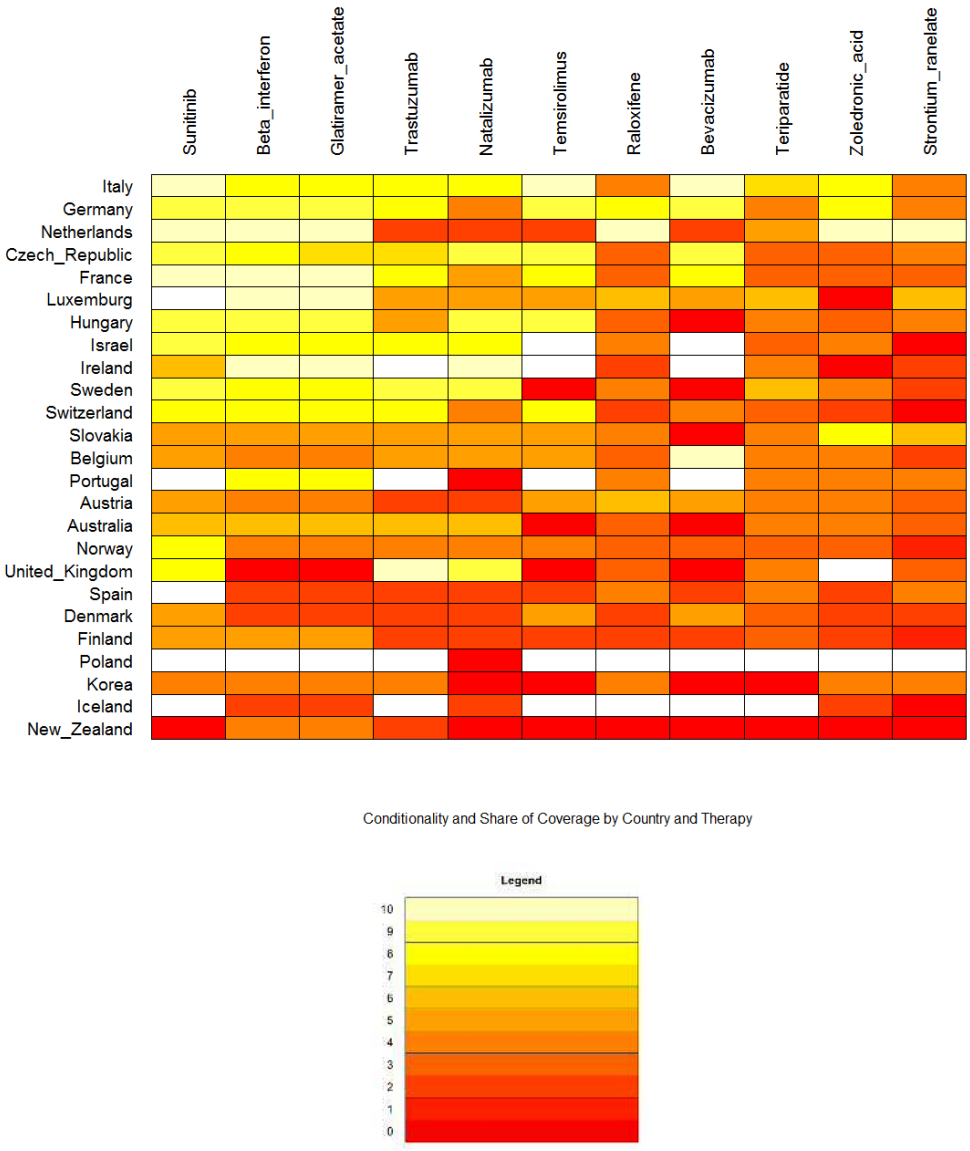
1. Introduction

One of the many challenges facing health care systems in developed democracies is the increase in costs for high-tech drugs. Hundreds of patented and often very expensive new pharmaceuticals enter markets every year and are demanded by patients. However, public health care systems in most countries have difficulties financing all drugs that are demanded, and perhaps also required, by patients. Where no limits to the public coverage for high-tech drugs are set, costs are likely to burst budgets for pharmaceuticals and produce opportunity costs elsewhere inside and outside the health care system. In 2008 pharmaceuticals accounted for 16.9 percent of total health expenses in OECD countries, compared to an average of 13.5 percent in 1980 (OECD 2011). Among the most frequently demanded and most expensive new pharmaceuticals are drugs for malignant, chronic and age-related diseases.

We have gathered data on the public financing of 11 controversial drugs in 25 OECD-countries, considering both the conditions under which these drugs are available to patients and the share to which their price is covered by the public systems (leaving the difference between reimbursed cost and the actual price to be paid out-of-pocket by the patient). We look at three groups of drugs: advanced drugs for malignant diseases (breast and kidney cancer), drugs for multiple sclerosis, a chronic degenerative disease that is difficult to treat, and drugs for osteoporosis, a serious, although non-fatal disease that typically affects post-menopausal women.

Figure 1 displays the extremely heterogeneous picture to be found. It plots the extent to which these controversial drugs are publicly financed in a color grid structured by countries (vertical axis) and pharmaceuticals (horizontal axis). To ease interpretation, countries are sorted in descending order of coverage from top to bottom and pharmaceuticals are sorted in descending order of coverage from left to right. To begin with, some drugs are more likely to be covered than others. This is hardly surprising, given differences in price and effectiveness. Moreover, however, different countries cover different drugs and some countries are generally more ‘generous’ than others, i.e. they cover more controversial drugs. According to our estimates, Italy is the most generous country in our sample, whereas New Zealand is the least generous.

Figure 1: Generosity of coverage by country and pharmaceutical



Note: Blank cells indicate missing values. Countries are sorted in descending order of generosity from top to bottom; pharmaceuticals are sorted in descending order of generosity from left to right. More specifically, countries and drugs are ordered by estimated intercepts obtained from an "empty" non-nested multilevel model. The model estimated is structurally similar to the specification described in section 4. The model is, however, "empty" in the sense of containing no explanatory variables.

What explains those differences in public generosity where the coverage of high-tech drugs is concerned? Is generosity merely a function of economic wealth, as expressed in the gross national product (GNP)? Does it depend on a society's across-the-board generosity for health care, as expressed in public health expenses per capita? Do we need to consider institutional factors like the history and structure of the respective health care system? Finally, given that decisions on the coverage of new drugs are in nearly all OECD countries delegated to specialized bodies or agencies (cf. Landwehr & Böhm 2011), we need to address the structure of the decision-making process as a potential explanatory factor for differences in generosity.

This paper seeks to answer these questions by way of a multilevel analysis of coverage decisions in the OECD world. Our analysis is based on OECD health data as well as a novel dataset for both coverage decisions and decision-making processes in 25 OECD-countries containing data we collected in a research project in 2010. We come to the surprising result that neither economic wealth nor expense levels significantly affect generosity. Instead, explanations need to consider the institutional set-up of the health care system and, more importantly, the structure of processes in which decisions whether to publicly finance a drug or not are made. Put briefly, the explanations for generosity in the public financing of high-tech drugs are not economic, but political ones. We conclude that if institutional design affects financing decisions, institutional design is a matter not only of procedural, but also of distributive justice.

2. Coverage decisions in health care: existing research, theories and hypotheses

Coverage decisions on expensive pharmaceuticals are a central aspect of a more general problem: the need to set limits to health care spending in face of increasing costs and decreasing public revenues. Health care priority-setting, or, put more crudely, health care rationing, has in recent decades increasingly become a hot issue in public and academic debates. The 1980ies and early 90ies were characterized by a search for something like "objectively fair" allocation principles from which single coverage decisions could simply be derived. However, empirical experience with participatory and expert bodies

has shown that principles of distributive justice are way too abstract to guide concrete coverage decisions and that consensus on allocation criteria remains out of reach. The debate has hence taken a "procedural turn" (Holm 2000), thus focusing on desiderata for a fair decision-making process rather than on substantial coverage decisions. A seminal contribution in this regard has been Norman Daniels and James E. Sabin's model "accountability for reasonableness", in which the authors outline conditions bodies charged with coverage decisions need to fulfill in order to guarantee procedural justice (Daniels and Sabin 2002).

Until recently, most research on health-care priority setting focused on the fundamental ethical questions involved, using the case as an example to discuss implications of different theories of justice or outlining ideal decision-making procedures. In the last few years, however, more empirical literature has emerged. A major European project has analyzed the content of health baskets in numerous countries (Schulenburg and Blanke 2004), and several studies have compared decision-making processes in smaller sets of countries (e.g. Ham and Robert 2003; Sabik and Lie 2008; Landwehr and Böhm 2011). None of these contributions, however, has tried to account for differences in coverage decisions and resulting health baskets, and none has studied effects of decision-making procedures on resulting decisions. This is the research gap this paper seeks to address by asking which variables can account for differences in coverage decisions and by explicitly taking into account the effect of institutional factors (i.e. properties of decision-making processes) on resulting decisions.

What might explain differences in generosity in the public funding of high tech drugs? The first explanations that suggest themselves almost self-evidently are the economic wealth of a country and its general expenditure levels in the health care sector. Wealthier countries, so the apparently obvious explanation, can afford higher public health expenses and can cover more, and more expensive, drugs. Similarly, a society that can afford to spend a higher total amount per capita on health care than another and decides to do so within a public health care system seems likely to cover more, and more expensive, drugs. In a first step, we thus assess the effects of economic wealth and health care expenditure on coverage decisions for our eleven controversial high-tech drugs for

the treatment of four different diseases. The hypotheses to be tested here are straightforward:

- H1: The wealthier a country is (measured by GNP p.c.), the more 'generous' is the public funding of controversial drugs (i.e. the less restrictive are conditions and the higher the share that is publicly funded).
- H2: The higher *overall* spending on health care (p.c.) is, the more 'generous' is the public funding of controversial drugs.
- H3: The higher *public* spending on health care (p.c.) is, the more 'generous' is the public funding of controversial drugs.

In a second step, we assess whether the broader institutional context within which coverage decisions are made affects the extent to and conditions under which controversial high-tech drugs will be publicly financed. To begin with, the health care system as whole could clearly be a significant contextual variable. Health care systems are typically grouped into one of three categories: state health systems, social insurance systems and private systems. The traditional criterion for the assignment of systems to one of the categories is the way a system is financed. State systems are (predominantly) tax-funded, social insurance systems are (predominantly) funded through insurance contributions and sources of private systems (predominantly) arise from private (for profit and non-for-profit) insurances or from out-of pocket payments. On the basis of welfare state theory and existing research on health care systems, it is straightforward to hypothesize that in private health care systems, coverage of controversial drugs within the public part of the health care system is unlikely, as the publicly funded systems like, in the US, Medicaid, cover only basic treatment. As our sample does not include any private systems, we concentrate on differences between state health care systems and social insurance systems. Regarding the effects of the health care system (state vs. social insurance) on generosity in coverage decisions, we theorize that in social insurance systems, claims to treatment are viewed as a kind of 'property rights' obtained through previous contributions whereas in state systems, claims are directed at the community as such and more clearly dependant upon fiscal limits. Our hypothesis is thus:

- H4: In tax-financed and state-regulated health care systems (state systems) public funding of controversial drugs is less generous.

A second potentially important contextual variable is the way the *default* is set. The default is the outcome that is effected if no decision is taken. Its significance for actor constellations and resulting decisions has impressively been pointed out by Elinor Ostrom (1986). With a positive default, all pharmaceuticals are reimbursed until a contrary decision is made. In this case merely decisions to exclude a drug need to be taken. More common are negative defaults, where pharmaceuticals are funded after a positive decision only. We expect the default to have significant predictive power for generosity with a positive default rendering regulations more generous. This expectation is in accordance with game-theoretic work that emphasizes the crucial role default regulations play in shaping the power of individual actors with diverging preferences in individual decision problems. Roughly speaking, in case of a positive default, an actor preferring full coverage just needs to block any decision to the contrary.[1] We thus expect:

- H5: Where the default is negative, public funding of controversial drugs is less generous.

Most importantly, however, we also expect properties of the decision-making process itself to affect coverage decisions. All OECD-countries have set up specialized bodies to deal with the challenge of health care priority-setting, but have chosen very different institutional designs for these (see Landwehr and Böhm 2011). In order to be able to compare the various and often complex processes we suggest a set of categories that grasp the main institutional characteristics we theorize to have effects on resulting decisions (cf. Landwehr and Böhm 2011).

The first categories are the degree of *delegation* of decision-making power from government to non-governmental bodies and the *independence* of those from government and regulatees influences the outcome of decisions. Our theoretical expectation is that delegation and independence render public funding less generous: delegation to independent bodies constitutes a strategy of ‘depoliticization’ and possibly also ‘blame avoidance’ (Weaver 1986) with which elected governments evade responsibility for

unpopular decisions. Opportunistic politicians have incentives to *not* restrict the funding of certain drugs if this is expected to prove electorally costly. The general expectation from the welfare state literature is that social policy retrenchment is indeed associated with negative consequences at the ballot box; especially to the extent that losses are concentrated and potential gains diffuse (Giger 2011: 19). Where independent bodies enjoy much leeway, electoral costs are thus likely to be less relevant for the final decision. With more delegation and independence, less generous decisions are, therefore, to be expected.

Furthermore, we suppose that the *inclusiveness* of the committee, namely its composition and size, affects decision-making. Transaction cost and negotiation theory (most seminal: Coase 1937) have pointed out that the larger and more heterogeneous, and thus more inclusive, a committee is, the more costly and difficult, and thus eventually unlikely, decisions will become. A higher number of different actors (or actor groups) means that a higher number of divergent interests and points of view have to be coordinated, which affects the probability and content of decisions. However, it must also be noted that heterogeneity in the composition of a forum enlarges the argument pool actors can draw on and might thus improve the quality and acceptability of decisions. As to the degree of generosity to be expected from composition and size, our weak theoretical prior is that inclusiveness generally leads to more generous financing schemes. Inclusiveness is to a considerable degree determined by whether patients, lay persons and industry representatives are involved. We expect that these actors push the decision towards more generous public funding.

Apart from the composition of a forum and the number of members, the *decision rule* applied is a central aspect of its inclusiveness. Consensus requirements obviously render decisions more difficult, while the availability of majority decisions (and hierarchical decisions even more so) reduces transaction costs and thus increases the probability that a decision will be reached at all. How this affects generosity should be contingent on the default regulation in place. When the default is negative (as is true for 23 out of the 25 countries under investigation), more inclusive decision rules should lead to less public funding. Considered from a game-theoretic perspective, under a consensus rule and a negative default the final decision should reflect the preferences of the actor

who is most strongly opposed to public funding as this actor can simply block any decision that is more generous than his ideal point.

Finally, the *transparency* of the decision-making process might have an impact on the generosity of decisions. Transparency is marked as a central desideratum in the normative debate on procedural justice of distributive decisions because it is believed to increase accountability (Daniels/Sabin 2002). However, we also expect the degree to which a decision-making process is publicly accessible and transparent to have effects on resulting decisions, although the direction of these effects is less clear. On the one hand, publicity and transparency increase opportunities for public scrutiny and may thus make unpopular decisions (i.e. ones not to cover a drug) more difficult. On the other hand, publicity and transparency may provide decision-makers with opportunities to justify tough decisions, thus eventually facilitating these. To summarize, the following hypotheses on the effects of institutional characteristics of the decision-making process follow from our theoretical considerations:

- H6: Where the degree of delegation and independence of an appointed body are high, public funding of controversial drugs is less generous.
- H7: The more inclusive an appointed body, the more generous is the public funding of controversial drugs.
- H8: The higher the majority requirements for decisions (maximum: consensus), the less generous is the public funding of controversial drugs.
- H9a: The higher the transparency of the decision-making process, the *more* generous is the public funding of controversial drugs.
- H9b: The higher the transparency of the decision-making process, the *less* generous is the public funding of controversial drugs.

3. Study design and data

Our study addresses coverage decisions for high-tech drugs in 25 developed democracies. We consider the OECD member countries as of 2009, including Israel which acceded 2010 and excluding Mexico, Turkey and Greece for which the quality of available data is too poor, as well as the United States and Canada, where the decentralized character of

the public health care system does not allow general claims on the public coverage of single drugs.

Concerning our dependent variable, we have chosen eleven pharmaceuticals for the treatment of four different diseases. This choice is the result of a two-stage selection process that was guided by the search for variance in decision-outcome and the wish to include drugs for different kinds of diseases. In a first step, we scanned reimbursement decisions and assessment reports for a smaller country sample to find innovative drugs for which reimbursement was contested due to controversial efficiency and for which coverage decisions varied between countries. In a second step, we selected condition-treatment pairs in order to represent a range of different diseases, including a "normal" age-related medical condition (osteoporosis), a severe but not terminal illness (multiple sclerosis) and two types of cancers of which one is very common but treatable (breast cancer) while the other is rather rare but highly lethal (renal cell carcinoma). For three of the conditions (osteoporosis, multiple sclerosis, renal cell carcinoma) for which treatments were considered, our sample contains several controversial treatments, which may in some cases substitute for one another. In order to get full information on generosity, we thus considered all available and controversial treatments for a given condition: a country that does not fund drug A, but drugs C and D for a given condition may be viewed as more generous than one that funds A, but not C and D.

Zoledronic acid, teriparatide, strontium ranelate, and raloxifene are used for the treatment of *osteoporosis*. All four drugs are very much more expensive than standard therapy without securely proven better effectiveness. Beta-interferon, glatiramer acetate and natalizumab are immune-modulating substances for the treatment of relapse-remitting *multiple sclerosis* (MS). Trastazumab (Herceptin) is a monoclonal antibody for the adjuvant therapy of a certain type of non-metastatic *breast cancer* that is given after initial treatment (e.g. chemo therapy) to improve long-term prognosis. In the case of trastazumab, the required duration of treatment is contested. Sunitinib, temsirolimus and bevacizumab have market authorization in the countries studied for the first-line treatment of metastatic *renal cell carcinoma*.

To measure the generosity of public reimbursement, we have developed two indices that contain information about the *conditions of coverage* and the *share of*

coverage pertaining to the individual country-drug cases which we then combine into one summary measure of generosity. To be able to compare conditionality of reimbursement between countries, we have first collected data on constraints financing of these pharmaceuticals is subject to and which are not named in its market authorization. In a second step we compared single conditions and classified them according to their potential to restrict coverage. We then categorized each pharmaceutical according to restrictions on coverage into five levels of conditionality, ranging from unconditionally covered to not covered at all (see appendix for a description of the index and conditions categories). The resulting *conditions of coverage* index ranges from zero to one, with zero indicating "not covered" and one reflecting "unconditionally covered." The *share of coverage* index displays how much of drug costs have to be borne by the patient. If the particular drug is not covered, the index takes the value zero, and it takes the value one if it is fully reimbursed with no co-payments for the patient. In between, we distinguish four different levels of co-payment (for a detailed description of the index values and the calculation of co-payment, see appendix).

To reflect the fact that the generosity of public coverage depends clearly on both aspects, we combine the two indices into one overall *generosity*-index.[2] In our view, the most sensible way to combine this information is by computing the product of the two individual indices with equal weights attached to the two components. A multiplicative index has the advantage that, in contrast to an additive index, it does not allow for full compensation (see Nardo et al. 2005: 79), i.e. it is not possible for a country to compensate for very strict conditionality restrictions with a high share of coverage. We rescale the resulting generosity-index such that it takes a maximum value of ten (when pharmaceuticals are unconditionally and fully covered) and a minimum value of zero (when a drug is not covered at all) by multiplying the product of the two sub-indices with the factor ten in order to facilitate the description in the empirical analyses.[3]

Regarding the independent variables identified in section 2, we have drawn data on GNP per capita and public health expenses as well as on the financing of health care systems (all data are for 2008) from OECD health data 2011 (OECD 2011). In order to operationalize the properties of the decision-making processes, we have developed indices which translate qualitative information on the particular institutional

characteristics into numerical values between zero and one.[4] Merely information on the default is included as dummy variable. Drawing on an independence index for regulatory agencies suggested by Gilardi (2002), we have established a delegation and independence index which considers the members' status, the body's financial and organizational autonomy, the competences of the committee and which asks who takes the final and binding decision. The inclusiveness of the involved committee is mapped by an index which contains information about the members, the number of members and involved stakeholders. Due to its relative importance (see above section 2), we have not included the decision-rule into the inclusiveness index but have given it an own value. And last, we have used information on the public availability of proceedings, meetings and reports to build the transparency index. Details on the construction of the indices can be found in the online appendix accompanying this article where we also list the data on the independent variables and their pairwise correlations with each other.

4. Statistical Model

We explore the predictive power of the factors discussed above within a non-nested (or "cross-classified") multilevel model where individual regulations are simultaneously nested within countries as well as pharmaceuticals. Given that our observations are structured along these two dimensions the natural way of analyzing this data is to directly specify this grouping structure in the statistical model. In contrast to usual multilevel applications in political science that employ strictly hierarchical models (see Steenbergen/Jones 2002) our grouping structure is non-nested. In our data each individual observation simultaneously relates to a specific pharmaceutical *and* a specific country. In addition to the level of the individual observations (level 1), we thus have two group levels (level 2) that are not hierarchically related to each other. Hence, we specify distributions for the individual level, for the country-intercepts and the therapy-intercepts. Formalizing this discussion, our statistical model takes the following form (our notation builds on Gelman/Hill 2007):

$$y_i = \alpha + \mu_j + \delta_k + \varepsilon_i \text{ for } i = 1, \ldots, N$$

$\mu_j \sim N(\beta X_j, \sigma_\mu^2)$ for $j = 1, \ldots, J$

$\delta_k \sim N(\omega Z_k, \sigma_\delta^2)$ for $k = 1, \ldots, K$

In the individual level model (first column), the generosity of the regulation (y_i) is regressed on an overall intercept (α), on (j) country-specific intercepts (μ) and on (k) drug-specific intercepts (δ). The second column specifies a distribution for the country intercepts; it is this part of the model we are substantially interested in. We assume a normal distribution with variance (σ_μ^2) estimated from the data and a mean that is specified as the product of country specific vectors of predictors (X_j) and a vector of corresponding coefficients (β) to be estimated from the data. In the matrix of country-level predictors we include the explanatory factors discussed above. The third column specifies a distribution for the pharmaceutical intercepts. Again, we assume a normal distribution with variance (σ_δ^2) to be estimated from the data. Rather than just assuming a mean of zero, we directly model the fact that the eleven pharmaceuticals relate to four different diseases. In order do to that, we specify the mean of the distribution to be the product of drug-specific vectors of predictors (Z_k) and a vector of corresponding coefficients (ω). The columns of the matrix of drug level predictors are dummy variables for three of the four diseases (with the remaining one building the reference category). The vector ω is thus of length three and contains the coefficients for these three dummy variables.

We estimate this multilevel model via Winbugs (version 1.4.3), a statistical software that allows for Bayesian analysis using Markov Chain Monte Carlo (MCMC)-methods. In doing so, we follow the advice of Gelman and Hill (2007: chapter 16) who advocate the use of Bayesian MCMC-methods in fitting multilevel models in case of more complex grouping structures like non-nested ones. More generally, Bayesian estimation tends to produce more accurate and conservative results (than alternative Maximum Likelihood estimation) in cases where the number of level-two units is small (e.g. Stegmueller 2011). We use non-informative priors for coefficients and variances, again following the usual practice as presented in Gelman/Hill (2007). Convergence is checked via the potential scale reduction factor $\hat{R}$ that assesses convergences via the

mixture of different chains (see Gelman/Hill 2007: 358). All reported results are based on results from a sufficiently large number of iterations such that $\hat{R} < 1.1$.

In the result section, we present posterior means and lower and upper bounds of the corresponding 90% credible intervals for our parameters of interest. For readers more familiar with standard regression approaches, we note that these can be interpreted analogously to estimated coefficients and corresponding 90% confidence intervals from standard regression outputs. Likewise, a coefficient can be interpreted as being statistically significantly different from zero with $p<0.10$ if zero is not contained in the 90% credible interval. We present R^2-measures of explained variance at the different levels (data level, country level and pharmaceutical level) following Gelman and Pardoe (2006). These can be interpreted analogously to the classical (adjusted) R^2.

5. Results

In the estimation of our models, we proceeded in an explorative and stepwise fashion given the limited prior knowledge on our subject matter and to save degrees of freedom. This section presents the empirical findings. It follows the structure laid out above in that we first consider the economic wealth and general expenditure levels in the health care sector. We then turn to the institutional context within which coverage decisions are made. Finally, institutional parameters of decision-making processes and involved bodies are taken into account. Our approach is theory-guided in that we move from less proximate to more proximate factors in a manner that allows us to eliminate potential explanatory factors in turn. This way it becomes possible to estimate reasonable reduced models that do not contain all possible explanatory factors at once; such a latter "garbage-can" (Achen 2005) approach would be unwise given the limited information in our data.

Table 1 investigates into the predictive power of the first group of explanatory variables. Overall it presents the results from four different specifications that differ (only) with respect to the country-level predictors that are included. In the first three models only one predictor is included in turn. The first model introduces the Gross National Product per capita in thousand US$, the second model total health expenditure per capita in thousand US$ and the third model public health expenditure per capita,

again, in thousand US$. These variables are strongly correlated to each other such that it would be problematic in terms of multicollinearity to include them in one model at once. For none of these three variables, we do observe any noteworthy association with our dependent variable. The posterior means indicate small effects in all cases: For example, a move from the poorest country in our sample to the richest one corresponds to an expected increase in generosity of about only 0.5. Further as indicated by the credible intervals, in none of the case can we have any confidence that the effects are different from zero. The low predictive power of the three predictors is also evident by the poor model fit as indicated by the measure of explained variance at the country level: It is negative across these three models meaning that the estimated error variance is larger than the estimated variance of the country intercepts in these models.

Table 1: Wealth, health expenditure levels and generosity of public coverage of high-tech pharmaceuticals

	(1)	(2)	(3)	(4)
Variables country level				
GNP p.c. (in thousand)	0.007			0.010
	[-0.025; 0.040]			[-0.026; 0.046]
Total health expenditure p.c. (in thousand)		0.108		
		[-0.215; 0.438]		
Public health expenditure p.c. (in thousand)			-0.011	
			[-0.460; 0.407]	
Total health expenditure in % of GDP				0.149
				[-0.371; 0.691]
Public health expenditure in % of total health exp.				-0.030
				[-0.117; 0.053]
Variables therapy level				
Breast cancer	1.77	1.81	1.81	1.79
	[-0.42; 3.90]	[-0.66; 4.16]	[-0.48; 4.13]	[-0.59; 4.14]
Multiple sclerosis	2.14	2.12	2.10	2.12
	[0.53; 3.70]	[0.56; 3.70]	[0.51; 3.55]	[0.52; 3.65]
Renal cell carcinoma	1.50	1.50	1.47	1.48
	[-0.05; 3.12]	[-0.05; 3.05]	[-0.17; 3.05]	[-0.16; 3.12]
Standard deviation				
country level intercepts (σ_μ)	1.75	1.75	1.73	1.83
	[1.26; 2.34]	[1.23; 2.29]	[1.26; 2.31]	[1.32; 2.48]
pharmaceutical level intercepts(σ_δ)	1.09	1.11	1.10	1.09
	[0.54; 2.01]	[0.53; 2.01]	[0.53; 1.99]	[0.54; 1.98]
Explained variance				
overall	0.42	0.42	0.42	0.42
country level	-0.03	-0.02	-0.04	-0.09
pharmaceutical level	0.36	0.34	0.35	0.35
Number				
of observations	247	247	247	247
of countries	25	25	25	25
of pharmaceuticals	11	11	11	11

Bayesian non-nested multilevel models; dependent variable is the generosity of public coverage of pharmaceuticals; listed are posterior means for coefficients with 90% credible intervals given in brackets below the coefficients (i.e. values for the 5th and 95th percentile of the posterior distribution); constant and country and pharmaceutical level intercepts not shown.

Model 4 investigates into the question whether any visible pattern emerges when accounting for wealth, total and public health expenditure simultaneously. In order to do that we look at total health expenditure in percent of GDP and public health expenditure in percent of total health expenditure, i.e. we reconstruct these measures in such a way that collinearity is reduced. Our conclusions remain unchanged: To our surprise, the data do not support the hypotheses, that more wealthy states are more generous with regard to

the public financing of high-tech drugs (H1); nor is there any evidence to support the seemingly obvious conclusion that the generosity in the public financing of high-tech drugs consistently corresponds to general expenditure levels for overall (H2) or public health care (H3). As all the considered variables do not seem to be consequential for generosity, we do not include them in the following models.[5]

Table 2, thus, concentrates on the remaining two groups of explanatory variables. Model 5 introduces two variables that relate to the broader institutional context in which coverage decisions are made: The type of the health care system, i.e. whether it is predominantly societally or state funded, and a dummy variable that captures whether the default is positive or negative. Both variables seem to be clearly associated with generosity in the expected direction. The point estimates for the posterior means indicate that generosity is higher by 2.1 in countries with a positive default and 1.4 in countries with a societal health care system. These effects correspond to substantial differences in generosity given the range from zero to ten. Moreover, the 90% credible intervals do not contain zero in both cases such that we can reject the nil hypotheses with reasonable confidence. The model explains 28% of the variance across countries; a substantial improvement over the previous models. Initial support for the hypotheses that associate state funded health care systems with less generosity (H4) and positive defaults with more generosity is thus obtained (H5). Given their predictive power we keep these variables in the subsequent models that introduce the variables that relate to institutional features of decision-making processes and delegative bodies.

Table 2: Health care system structure, delegative institutions and generosity of public coverage of high-tech pharmaceuticals

	(5)	(6)	(7)	(8)
Variables country level				
Health care system	1.39	0.84	0.80	
societal	[0.32; 2.42]	[-0.46; 2.16]	[-0.40; 1.92]	
Positive default	2.13	2.37	2.34	2.45
	[0.13; 4.00]	[0.28; 4.48]	[0.22; 4.38]	[0.34; 4.57]
Delegation and		-1.95	-2.30	-3.10
independence		[-4.62.; 0.87]	[-4.94; 0.15]	[-5.49; -0.49]
Transparency		1.36	1.45	1.75
		[-0.50; 3.12]	[-0.23; 3.12]	[0.24; 3.40]
Inclusiveness		0.29		
		[-2.52; 3.05]		
Decision rule		-1.98	-2.08	-2.27
		[-3.95; -0.06]	[-3.99; -0.26]	[-4.05; -0.33]
Variables therapy level				
Breast cancer	1.81	1.73	1.77	1.80
	[-0.64; 4.18]	[-0.69; 3.96]	[-0.73; 4.26]	[-0.34; 4.08]
Multiple sclerosis	2.13	2.15	2.12	2.20
	[0.55; 3.74]	[0.52; 3.78]	[0.49; 3.76]	[0.68; 3.90]
Renal cell carcinoma	1.48	1.44	1.47	1.51
	[-0.23; 3.15]	[-0.06; 2.98]	[-0.13; 3.10]	[-0.01; 3.05]
Standard deviation				
country level intercepts	1.43	1.42	1.36	1.40
(σ_μ)	[1.00; 1.97]	[0.93; 2.04]	[0.90; 1.94]	[0.99; 1.97]
pharmaceutical level	1.12	1.11	1.12	1.08
intercepts(σ_δ)	[0.53; 2.04]	[0.53; 1.96]	[0.55; 2.02]	[0.47; 2.00]
Explained variance				
data level	0.42	0.42	0.42	0.42
country level	0.28	0.34	0.38	0.35
pharmaceutical level	0.32	0.35	0.32	0.36
Number				
of observations	247	247	247	247
of countries	25	25	25	25
of pharmaceuticals	11	11	11	11

Bayesian non-nested multilevel models; dependent variable is the generosity of public coverage of pharmaceuticals; listed are posterior means for coefficients with 90% credible intervals given in brackets below the coefficients (i.e. values for the 5th and 95th percentile of the posterior distribution); constant and country and pharmaceutical level intercepts not shown.

All four concerned variables are introduced in model 6. With the introduction of these variables the health care type dummy loses in predictive power: The mean coefficient is cut in half and the 90% credible interval now also contains negative values. The results for the positive default remain largely unchanged. Coming to the newly introduced variables, we observe the clearest effect for the decision rule. Where decision

rules are more inclusive stipulating a consensus among the participant actors, generosity tends to be lower by about 2 points (as compared to hierarchical and 1.5 point as compared to majority decision rules) and the effect is reliably estimated to be negative. This supports H8. In contrast, inclusiveness is not related to generosity: The posterior distribution is wide with a mean close to zero. Thus, we do not obtain support for the proposition that inclusive committees will produce more generous regulations for the public financing of high-tech drugs (H7). Our results remain inconclusive with regard to delegation and independence as well as transparency. The posterior mean points to a substantial negative effect of delegation and independence which suggests that delegation and independence facilitate rationing decisions and more restrictive regulations as expected by H6. However, the coefficient for delegation and independence is imprecisely estimated with a large credible interval that is mainly in the negative range but also contains values above zero. A similar pattern emerges for transparency: The point estimate suggest a substantial positive effect with more transparency being associated with more generous regulations (as suggested by H9a). But again, the coefficient is imprecisely estimated and the 90% credible interval also marginally contains negative values. The data thus point for both delegation and independence and transparency (as far as H9a is concerned) in the expected direction, but the confidence in these associations remains limited. Model 7 eliminates inclusiveness from the set of country-level predictors. The estimation results for the remaining variables stay similar in comparison to model 6. As indicated by the explained variance at the country level in both models 6 and 7, the introduction of the institutional features of decision-making processes and delegative bodies further helps to improve the predictive accuracy of the model.

Finally, model 8 excludes the health care dummy. While this is partly problematic given the predictive value of this variable (especially in model 1), the exclusion is based on the rationale that some of the features of the delegative institutions are to some extent related to the overall type of the health care system for reasons elaborated upon elsewhere (AUTHORS 2011). For this reason, it might be hard to statistically reliably distinguish effects of the health care system from those of these institutional features. Therefore, model 8 explores how the exclusion of the health care type changes the estimation results. The results for the default regulation and decision rule remain roughly unchanged

with substantial and reliable associations in the theoretically expected directions. In case of delegation and independence as well as transparency, we also observe only slight modifications which however would lead to different conclusions from the perspective of statistical significance testing. Both posterior means are larger in absolute value as compared to models 6 and 7. What is more, the credible intervals do contain only negative values for delegation and independence and only positive ones for transparency. Accordingly, the nil hypothesis could be rejected with $p<0.10$ in both cases. Overall, there are considerable hints that these two variables are reliably associated with generosity in the respective directions. However, our results also make clear that it is hard to definitively establish this and to statistically distinguish this association from effects of the health care system type on the basis of the data at hand.

6. Discussion

The first and rather surprising finding of our study is the fact that neither the wealth of a country nor the level of public health expenditure affects the extent to which high-tech drugs are covered within public systems. Clearly, our study is limited to explicit regulations on particularly controversial drugs and does not assess whether access to these is limited through mechanisms of implicit rationing. Governments may choose to avoid explicit decisions, which are politically explosive, and instead try to limit expenses by way of tight budgets for hospitals and doctors. Nonetheless, our results indicate that political and societal attitudes towards the coverage of high-tech drugs are at least not predominantly driven by wealth and expenditure patterns.

Results have also confirmed our hypotheses that the financing of the health care system – through taxes or social insurance contributions – affects its generosity: In countries with social insurance systems, high-tech drugs are significantly more likely to be covered. As noted in section 2, this difference may be accounted for by the different quality of claims to health services. In social insurance systems, claims are directed at health funds and backed by previous contributions, whereas in state systems, claims are directed at the community of tax-payers and more likely to be made dependant upon the fiscal situation. This difference between types of health care systems also accounts for the

fact that the discussion on limit-setting and rationing in health care has reached the conservative welfare states of continental Europe decades after it came up in the Anglo-American and Scandinavian countries. Only quite recently, the costs associated with the public coverage of expensive, patent-protected drugs have become an issue in continental Europe, and attempts at limit-setting are so far faltering in countries like Germany or France.

Given that the institutional properties of the health care system seem to have significant effects on the coverage of high tech drugs, it made sense to take a closer look at the decision-making process that regulate coverage decisions. Our results confirm the significance of the way the default is set, showing that a positive list for drugs reduces the probability that controversial products are funded. Only three countries in our sample do not apply positive lists: Germany, the UK and Ireland. Given the overall shortage of resources in the British system, the default for the UK is negative, as primary care trusts are unlikely to cover an expensive drug that has explicitly been recommended to not be founded by the National Institute of Health and Clinical Excellence (NICE), which issues technology appraisals for the UK. In Germany and Ireland, the positive default clearly seems to contribute to generosity. In case of the former, several attempts to introduce a positive list (and thus negative default) for drugs have failed, possibly due to pressure from the strong German pharmaceutical industry.

However, not only the default, but also properties of the appointed bodies charged with coverage decisions are associated with the generosity of resulting decisions. Finding that the decision rule is the most significant factor confirms rationales drawn from transaction cost and negotiation theory. It may seem surprising that consensus requirements render positive, and thus popular, decisions less likely, though. However, this may be accounted for by the fact that in the majority of our cases, the default is negative. This implies that the decision at stake, which is rendered difficult by consensus requirements, would be one to cover, rather than to exclude a service. For the other institutional properties of bodies charged with coverage decisions, the results are not robustly significant, but point into the hypothesized direction. More specifically, we obtain hints that delegation and independence are associated with less generous regulations, whereas transparency seems to be associated with more generosity. The

corresponding coefficients, however, are only statistically significant in one of the models.

Given the comparatively small number of countries, more robustly significant results were unlikely to occur. The confidence in the theoretical significance of our results has to remain limited to some extent. Definitive conclusions on all possible explanatory variables are ruled out. Nonetheless, our findings strongly suggest focusing on institutional rather than monetary factors when trying to explain countries' generosity in the funding of high-tech drugs. A deeper understanding of the causal relationships between institutional properties, especially those of the bodies charged with decision-making, would require in-depth case studies and qualitative analysis. This remains a clear desideratum for future research.

7. Conclusion

The wealth of a country and its level of public health care spending do not affect the extent to which innovative but extremely expensive drugs are funded within a public health care system. As previous research indicates, wealth and spending do not affect health outcomes, measured, for example, by healthy life expectancy, either (Kotzian 2009). We therefore come to the conclusion that 'generosity' in the public financing of advanced medical technologies is eventually a matter of societal values and interest constellations. These affect decisions, among other things, by being inscribed into the structure of the health care system as such as well as into decision-making processes and the institutional design of bodies appointed to take coverage decisions. Apparently, institutions do matter for the allocation of health care goods and services, and they matter more than wealth and fiscal constraints. Institutional design in the health care system should thus not only be addressed as a matter of procedural, but also as a matter of distributive justice.

Notes

[1] A further refinement of this argument is the hypothesis that there might be some interaction between the default regulation and the decision rule, specifically whether decisions are taken by consensus or some other rule. The limitations of our data with only two empirical cases of a positive default prevent us from specifying such an interaction in our statistical models.

[2] A similar logic has been regularly employed in the construction of generosity or decommodification indices of welfare state programs in the tradition of Esping-Andersen (1990) that combine information on "replacement rates", on the one hand, and "qualifying conditions", on the other (e.g. Scruggs/Allan 2006).

[3] The combined index thus also contains the missing values of both indices. Hence, missing values result from either missing information on the share of coverage or on the conditionality of coverage or both. In some cases (Ireland, Iceland) we did not include decisions that were taken by the precursor committee because institutional characteristics of the old committees differed widely from those of the new committee. In Poland, we had no access to the positive list and thus could only include information from secondary sources. The resulting index can take 17 unique values (of which 15 are observed in our data) and it seems legitimate to treat it as an interval-scaled measure in the empirical analyses thereby imposing a linearity assumption. Specifying an ordered logit model would not be feasible in terms of costs in degrees of freedom given our data and the 15 observed unique values in our dependent variable.

[4] All data on the institutional parameters of decision-making processes and the dependent variables was collected in an observation period in the first half 2010. Thus, neither coverage decisions nor changes in institutional design after June 2010 were considered.

[5] Nonetheless, we have checked this strategy by reintroducing (each variable in turn) the three explanatory variables from models 1 to 3 from table 1 into our preferred model 3 from table 2 to be discussed below. None of the wealth and expenditure variables emerges as substantially or statistically significant with these other explanatory variables included. Moreover, the results for the other variables remain reasonably similar.

References

Achen, C.H. (2005) 'Let's Put Garbage-Can Regressions and Garbage-Can Probits Where They Belong', *Conflict Management and Peace Sciences* 22(4): 327-339.

Coase, R. (1937) 'The Nature of the Firm', *Economica* 4(16): 386-405.

Daniels, N. and Sabin, J.E. (2002) *Setting Limits Fairly. Can We Learn to Share Medical Resources?* Oxford, Oxford University Press.

Esping-Andersen, G. (1990) *The Three Worlds of Welfare Capitalism*, Cambridge: Cambridge University Press.

Gelman, A. and Hill, J. (2007) *Data Analysis Using Regression and Mutlilevel/Hierarchical Models*, Cambridge: Cambridge University Press.

Gelman, A. and Pardoe, I. (2006) 'Bayesian Measures of Explained Variance and Pooling in Multilevel (Hierarchical) Models', *Technometrics* 48(2): 241-251.

Giger, N. (2011) *The Risk of Social Policy? The Electoral Consequences of Welfare State Retrenchment and Social Policy Performance in OECD Countries*, London/New York: Routledge.

Gilardi, F. (2002) 'Policy Credibility and Delegation to Independent Regulatory Agencies: A Comparative Empirical Analysis', *Journal of European Public Policy* 9(6): 873-893.

Ham, C. and Robert, G.B. (2003) *Reasonable Rationing: International Experience of Priority Setting in Health Care*, Philadelphia: Open University Press.

Holm, S. (2000) 'Developments in the Nordic Countries: Goodbye to the Simple Solutions', in Couter, A. and Ham, C. (eds.), *The Global Challenge of Health Care Rationing*, Buckingham/Philadelphia: Open University Press, pp. 29-37.

Kotzian, P. (2009): 'Productive Efficiency and Heterogeneity of Health Care Systems: Results of a Measurement for OECD Countries', *The Open Economics Journal* (2), 20-30.

Landwehr, C. and Böhm, K. (2011) Delegation and Institutional Design in Health Care Rationing, *Governance*, Volume 24, Issue 4, pages 665–688.

Nardo, M., Saisana, M., Saltelli, A. and Tarantola, S. (2005) *Tools for Composite Indicators Building. European Commission, Report EUR 21682 EN*. Joint Research Centre, Ispra, Italy.

OECD (2011) *OEDC Health Data 2011*, http://www.oecd.org/health/healthdata.

Ostrom, E. (1986) 'An Agenda for the Study of Institutions', *Public Choice* 48(1): 3-25.

Sabik, L.M. and Lie, R. K. (2008) 'Priority setting in health care: Lessons from the experiences of eight countries', *International Journal for Equity in Health* 7(4).

Schulenburg, J.-M. and Blanke, M. (2004) *Rationing Of Medical Services In Europe: An Empirical Study/A European Survey*. Amsterdam: IOS Press.

Scruggs, L. and James, A. (2006) 'Welfare-State Decommodification in 18 OECD Countries: A Replication and Revision', *Journal of European Social Policy* 16(1): 55-72.

Steenbergen, M.R. and Jones, B.S. (2002) 'Modeling Multilevel Data Structures', *American Journal of Political Science* 46(1): 218-237.

Stegmueller, D. (2011): *How Many Countries Do You Need for Multilevel Modeling?* Working Paper, University of Oxford.

Weaver, R.K. (1986) 'The Politics of Blame Avoidance', *Journal of Public Policy*, 6(4): 371-398.