UNIVERSITY OF WARSAW | FACULTY OF ECONOMIC SCIENCES

WORKING PAPERS
No. 10/2025 (473)

VALUING THEATER PERFORMANCES THROUGH BENEFIT TRANSFER: ACCURACY OF TRANSFERS OVER SPACE

EWA ZAWOJSKA
BARTOSZ JUSYPENKO
ALEKSANDRA WIŚNIEWSKA

WARSAW 2025

ISSN 2957-0506

WORKING PAPERS 10/2025 (473)

Valuing theater performances through benefit transfer: Accuracy of transfers over space

Ewa Zawojska, Bartosz Jusypenko, Aleksandra Wiśniewska*

Faculty of Economic Sciences, University of Warsaw

**Corresponding author: ewa.zawojska@uw.edu.pl*

Abstract: Understanding the value of cultural goods such as performing arts is essential for designing welfare-maximizing cultural policies. However, since these goods often generate public externalities, market data alone cannot capture their full value. While primary non-market valuation studies offer robust insights, they are resource-intensive. Benefit transfer (BT)—using value estimates from existing studies to assess unstudied contexts—offers a cost-effective alternative. This study examines the accuracy of BT in cultural economics, focusing on marginal values of theater performances. Using data from a discrete choice experiment (DCE) conducted in six Polish provinces, we perform inter-provincial transfers of marginal willingness-to-pay values for four types of theater performances. By comparing transferred values with primary estimates, we assess BT's validity and reliability. To our knowledge, this is the first BT study in cultural economics using DCE-derived marginal values. Our findings suggest that BT accuracy in culture aligns with results from more established fields, supporting the use of DCE-based estimates in cultural policy evaluation.

Keywords: accuracy, benefit transfer, discrete choice experiment, performing arts, reliability, validity

JEL codes: H43, Z11, Z18

Acknowledgements: Ewa Zawojska acknowledges the support of the Polish National Agency for Academic Exchange under the National Component of the Bekker Programme. Bartosz Jusypenko acknowledges the support of the National Science Centre of Poland (grant no. 2018/30/E/HS4/00388 and 2021/41/N/HS4/04321). The survey data was collected as a component of research grants funded by the National Science Centre of Poland (no. 2014/15/N/HS4/01328 and 2014/15/N/HS4/01441).

1 Introduction

Performing arts include all creative activities performed in front of an audience, such as concerts, opera, ballet, and theater. They can be classified as experienced goods, characterized by an individual demand that grows with increasing cultural consumption experience (Seaman 2006; Brito and Barros 2005). They provide a range of various benefits to people (Throsby and Withers 1979; Throsby 1990), including benefits arising from a direct contact with art (that is, use value) and benefits generated without a direct contact (that is, non-use value; for example, deriving a value from the existence of performing arts as a form of culture preservation). To assess the effectiveness of cultural policies aimed at provision of performing arts, it is necessary to know a monetarily-measured value of the benefits that performing arts bring to society. Comparisons of the money-metric measures of the benefits with costs (e.g., in benefit-cost analysis) help design cultural policies that can improve social welfare. While market transactions approximate the value of some goods based on the transaction prices,[1] performing arts are argued to generate positive public externalities (e.g., an educational role in society; Hansen 1997), which are not traded in markets. When goods generate benefits to those who do not "consume" them directly, market transactions fail to capture their total value. Then, non-market valuation approaches can be used. Primary non-market valuation studies typically require substantial resources, such as budget and time, and when such resource constraints exist, a benefit transfer (BT) method may be a feasible and cost-effective alternative (Johnston et al. 2018; Johnston et al. 2021). BT relies on using pre-existing value estimates from primary studies conducted at one or more sites or contexts (i.e., study sites) to calculate values at other sites or contexts (i.e., policy sites), for which limited or no primary data is available (Johnston et al. 2015b). While BT has been successfully employed in many fields, including health, recreation, and environmental economics (e.g., Brouwer and Navrud 2015; Loomis 2015), the method's application in cultural economics has been scarce (Lawton et al. 2021a). With this study, we intend to examine the potential for using BT in the area of culture and to help guide future applications of the method in this field. Our empirical case study used for the assessment of BT accuracy concerns theater performances.

While both, use and non-use benefits generated by performing arts are believed to contribute to the overall cultural enrichment of society, the latter typically constitute the primary argument to justify public financial support for performing arts, covering excessive production

[1] Valuation based on market prices relies on a simple premise that a rational individual does not want to pay for a good more than the good is worth to that individual.

costs associated with the labor-intensive nature of the industry (Throsby and Whiters 1979, Frey 1999).[2] The literature argues that performing arts generate positive public externalities and individuals who do not attend performing arts events may still derive benefits from the existence of performing arts (Throsby 1990, Hansen 1997). The externalities include, for example, preservation of culture, promotion of regional and national identity, as well as society's educational role. There also exist benefits to the art form itself, such as fostering innovation and supporting the development of local artists' creativity, which contribute to the evolution of art (Throsby 1990). The wide range of benefits generated by performing arts to society, which are not captured through market transactions (such as purchases of tickets for performances), indicate that non-market valuation tools need to be used to measure the total value that performing arts provide to society.

Non-market valuation can be conducted either through primary research or BT. Although primary research is generally preferred to estimate values, time and funding constraints faced in policy processes often lead to BT being the only feasible option (Johnston and Rosenberger 2010). For example, in the context of the Water Framework Directive of the European Union, the value estimates for benefit-cost analysis need to be obtained via cost-effective means, and BT is typically a much cheaper and faster method to estimate values than primary research (Hanley et al. 2006).

We aim to contribute to the underexplored question of the accuracy of BT when applied to experienced cultural goods, such as performing arts. We assess the accuracy through the framework proposed by Bishop and Boyle (2019), which focuses on two widely-used concepts within accuracy assessments: validity and reliability. A valid BT occurs when the transferred value estimate is not significantly different from the actual value of a given good; in other words, the transferred estimate is not biased. A reliable BT is associated with low transfer errors, which express the variance between the transferred and actual values. By examining the validity and reliability of BT for cultural goods, we intend to help guide future applied research whether (and in what conditions) this valuation technique can deliver accurate value measures for cultural policy purposes.

Empirical data to the study comes from a broad stated preference survey concerning the value of theater performances in Poland. The application of this valuation approach helps us

[2] In Europe, for example, theaters subsidized from public funds usually offer tickets for prices far below the level that would balance the costs of providing performances. Public subsidies often constitute more than half of public theaters' budgets (Zieba 2009, Zieba 2011), and, for example, in Poland, their share in theaters' budgets amounts, on average, to about 80% (Theatre Institute 2019).

measure the total value of theater performances including both use and non-use value (in contrast, for example, to revealed preference approaches based on people's actual choices, which allow for the assessment of use value only; Johnston et al. 2017). The survey is conducted on a sample from the general population of Poland, including both theatergoers and non-theatergoers. The stated preferences, necessary for the value estimation, are elicited in the survey through a discrete choice experiment (DCE). In the DCE, respondents are asked about their preferences towards various projects of extension of the scope of theater performances in their provinces. The projects of the theater scope extension are described by four types of performances: entertainment, drama, children's, and experimental. The approach allows for calculating the marginal value of extending theater performances for each of the types. The calculation is based on province-specific random parameter logit models. The models provide primary value estimates for provinces, which are treated as the actual values (that is, reference points) for the comparison with the transferred values.

For examining the BT accuracy, we focus on six provinces in Poland. We investigate all possible pairs and directions for BTs between these provinces, meaning that every province in our analysis serves as a study site, from which values are transferred to each of the other provinces, and as a policy site, to which values are transferred from each of the other provinces. Having the primary value estimate for a given province and the transferred value estimate (derived from the value estimate in another province), we are able to assess the validity and reliability of the inter-provincial BTs. We also attempt to identify factors and researcher's choices possibly enhancing the BT accuracy, by considering (i) a range of standard BT approaches, which include unadjusted unit value transfer, income-adjusted unit value transfer, and function transfer; (ii) different types of theater performances; and (iii) the degree of similarity between study and policy sites (i.e., different provinces). For each of the BT approaches, the analysis is based on 120 transfers of marginal values (six provinces times five other provinces, to which values are transferred, times four theater performance types) and 30 transfers of values capturing the benefits from maintaining the current state of theater performance in a province (six provinces times five other provinces, to which values are transferred).

Our examination of the BT accuracy is based on primary valuation research that was conducted in the Polish provinces in the same time period, with the use of the same data collection materials and methods, and with the application of the same data modeling techniques. We argue that this way, our investigation is a pure verification of BT accuracy for

transmission of information across locations (that is, over space), excluding confounding factors such as changes over time, variation in methodological practices, and differences in the definition of a good. As BTs are typically framed in terms of transmission of information over space (Johnston et al. 2018), this examination addresses this major dimension of differences in BT applications. The focus of this study on regions, according to administrative division, further excludes some confounding factors resulting from the choice of locations, such as currency discrepancies or, at the population level, large cultural differences or varying levels of political trust. These factors typically occur when transferring values from one country to another.

To the best of our knowledge, it is the first BT examination within cultural economics that is based on marginal value estimates from a DCE, focused on specific features of a proposed hypothetical program. Former studies focused on a total value of a program, with estimates most often derived from contingent valuation surveys.[3] DCEs typically involve a series of preference elicitation questions with policy projects defined by a set of attributes (such as theater performance types in our study). As a result, DCE data allows for estimating marginal values of the attributes. In turn, contingent valuation includes a non-attribute elicitation of preferences (e.g., a single yes-no choice), where a policy project is described as a whole. Contingent valuation is particularly useful to estimate values of goods inseparable into individual attributes.

The advantage of DCEs is that the marginal values can be used to evaluate several possible policy outcomes and design policies balancing different quantifiable goals. If DCE-based marginal value estimates are transferable, they can provide a flexible approach to valuation, allowing researchers to estimate the value of a cultural project at a policy site by selecting levels of attributes that are relevant to that site.[4] Furthermore, the relative number of DCE applications to contingent valuation applications appear to rise over recent years in non-market valuation literature (Mahieu et al. 2017; Hanley and Czajkowski 2019). We argue that this trend points to the importance of understanding accuracy of DCE-based BT, as more and more recent data available for BT may come from DCEs.

[3] We follow the distinction between DCEs and contingent valuation as described by the "Contemporary Guidance for Stated Preference Studies" (Johnston et al. 2017). However, we acknowledge that stated preference literature is not uniform with respect to the nomenclature. For example, some authors consider DCEs as a subset of contingent valuation when the latter is viewed as "a survey approach to place an economic value on a public good," "independent of any particular [preference] elicitation method" (Carson and Louviere 2011, p. 541-542).

[4] Naturally, the transferred marginal values need to be considered in the context of the entire primary valuation study, including the baseline level of the evaluated good, and to take into account possible diminishing marginal utility. We discuss this problem in more detail in the concluding section.

2 Benefit transfer as a valuation method

2.1 Usage of benefit transfer

BT may be understood as "the use of existing data or information on nonmarket values in settings other than where they were originally collected" (Rosenberger and Loomis 2017, p. 432). It is particularly helpful when resources like time, funding, or data are missing or insufficient to run primary research at a policy site, for which the value estimates are needed. Since such resource constraints are common within management and policy evaluations, BT is often then the only feasible approach for valuation. In the context of the US federal policy analysis, Newbold et al. (2018 p. 469) note that "it is impossible to conduct a prospective [benefit-cost analysis] without the use of at least some form of benefit (and cost) transfers." BT has been in frequent use in large-scale cost-benefit analyses in the US, EU, and elsewhere (e.g., Brouwer and Navrud 2015; Hanley et al. 2006; Loomis 2015; Rolfe et al. 2015b)—though, based on our thorough literature search, not yet much in the domain of culture.

The first applications of BT date back to the early 1970s and concern valuations of environmental goods, but the use of the method in cultural economics has a considerably shorter history. Rosenberger and Loomis (2017) refer to early examples of BT, including value estimates of recreation activities for assessing water-related projects published by the US Water Resources Council in 1973 and values of timber, forage, minerals, and water provided within the US Forest Service's Resource Planning Act in 1974. Further research has stimulated development of formal procedures of evaluating BT applications. The work by Freeman (1984) was likely the first one to propose conditions under which primary data could be transferable. It was followed by the 1992 section on BT in the journal of "Water Resources Research," where further recommendations for BT studies were formulated (e.g., Desvousges et al. 1992; Loomis 1992). Boyle and Bergstrom (1992) suggested that a valid BT requires identical study and policy sites, populations, and welfare measures. These strict conditions have been softened over time to provide pragmatic guidelines that take into account variations in research settings (Rolfe et al. 2015a). The book by Johnston et al. (2015a) is among the most recent comprehensive guides for this non-market valuation technique, but also developed in the context of transferring environmental and resource values.

As seen from the above largely abbreviated history of BT, the recommendations have been proposed and developed mostly in the area of environmental economics. With this paper, we want to contribute to understanding of the conditions for valid and reliable BTs in cultural economics.

We were able to identify nine BT studies in the area of cultural economics, including both published articles and reports, as well as one working paper and one PhD dissertation, as of September 2024.[5] We present an overview of the studies in Table 1. The small number of the identified BT studies within culture evidences that the application of the method in this area is yet in its infancy. For illustration of the method's use in other areas, we refer to a summary provided by Rosenberger (2015) in a chapter of the book on "Benefit Transfer of Environmental and Resource Values". Rosenberger (2015) analyzes BT studies that report transfer errors and identifies 38 such studies. 22 of them are within environmental and resource economics, 14 concern recreation, and two are about human health. This shows that the numbers of BT studies in environmental and resource economics, and recreation are much larger than the number for cultural economics. To keep the time span comparable, among our identified BT studies for culture, there are only four published before 2015 (i.e., the year of Rosenberger's publication).

[5] Although we made efforts to assure comprehensiveness of the literature search, we acknowledge that there might be studies that we were not able to identify.

Table 1. Review of BT applications in culture.

Author(s)	Publication type	BT study type	Country of the policy site(s)	Evaluated good(s) at the policy side(s)	Primary value estimates	BT(s) over space	BT(s) over time
Brown (2004)	PhD dissertation	Accuracy assessment	UK	Built heritage	Original: DCE	Across sites within the same country	No
eftec (2005)[a]	Report	Application	UK	Built heritage	Pre-existing: CV or TCM	From a site within the same country or from another country (Australia)	> 3 years
Tuan et al. (2009)	Journal publication	Accuracy assessment	Thailand, Vietnam	Built heritage	Pre-existing: CV	Across countries	6 months
Ulibarri and Ulibarri (2010)	Journal publication	Application	US	Archeological artifacts	Pre-existing: combined	From a site within the same country and from other countries (Australia, Canada)	> 3 years
Mourato et al. (2014)	Report	Accuracy assessment	UK, Germany, Sweden, Romania, Italy	Built heritage, museums	Original: CV	Across sites within the same country and across countries	No
Fujiwara et al. (2018)	Report	Accuracy assessment	UK	Museums	Original: CV	Across sites within the same country	No
Lawton et al. (2021b)	Report	Accuracy assessment	UK	Galleries, theaters	Original: CV	Across sites within the same country	No
Lawton et al. (2021c)[b]	Journal publication	Accuracy assessment	UK	Built heritage	Original: CV	Across sites within the same country	No
Wiśniewska et al. (2023)	Working paper	Accuracy assessment	Denmark, Poland	Theaters	Original: CV, DCE	Across countries	2 years

Notes: The literature search was completed in September 2024. Column "BT study type" informs on whether the study was solely an application of BT or an examination of BT accuracy. Column "Primary value estimates" informs about the source of the primary values that were used in a study: "pre-existing" means that the authors took value estimates elsewhere from the literature; and "original" means the authors used their own primary research to obtain estimates, which they next used for BT. The column also specifies the methods used in the primary valuation research: "CV" means contingent valuation, "TCM" – travel cost method. The column "BT(s) over time" reports the approximate time interval between the primary research and the policy scenario.

[a] One of the case studies included in this report is also described by Provins et al. (2008).

[b] Detailed information about the study is available in the research report Lawton et al. (2018).

The identified BT studies in culture concerned nearly exclusively built heritage assets (e.g., historic cities, temples, manor houses) and heritage institutions (museums). Only two studies valued art galleries or theaters. The BT findings were predominantly for the cultural sector in the UK–five out of the nine studies were located exclusively in the UK. The primary value estimates to the BTs were taken predominantly from stated preference contingent valuation studies, which were in some cases complemented with estimates derived from other methods, such as the travel cost method and contingent behavior statements. In one study (Brown 2004), the accuracy assessment is performed involving estimates derived from a DCE, though marginal values calculated for particular attributes are not transferred. Wiśniewska et al. (2023) use estimates from a DCE study, but only of the status quo option, to evaluate the current level of theater services. Different BT approaches were investigated, including unadjusted and adjusted unit value transfers, and function transfers.

Seven of the identified studies focused on the assessment of BT accuracy, whereas the remaining two were applications of the BT method. Six of the accuracy assessments first estimated the primary values based on their study-site data and next used the estimates to assess accuracy of transfers. This is also the approach employed in our investigation. The remaining accuracy assessment study used estimates from pre-existing valuation studies available in the literature.

All the identified studies performed transfers over space. However, the geographical distance between policy and study sites varied. In five studies, values were transferred across different sites (i.e., specific locations, where the considered assets are located) within one country, which could reduce potential sources of variation between the study and policy sites, for instance, due to limited cultural differences. Some inquiries included transfers between countries (e.g., Tuan et al. (2009) between Thailand and Vietnam, Mourato et al. (2014) between five European countries, and Wiśniewska et al. (2023) between Poland and Denmark). Ulibarri and Ulibarri (2010) employed pre-existing (true) value estimates from three different countries (Australia, Canada, US) to estimate the value at a policy site located in the US.

Transferring over time remains an unpopular practice in BT applications in culture due to the lack of preexisting original estimates. Estimates obtained long before the considered policy change were used only in two BT application studies in which the accuracy assessment was not performed, and thus nothing can be said about the impact of time distance on transfer precision. Tuan et al. (2009) used surveys conducted within a six-month interval in the same year, whereas Wiśniewska et al. (2023) used surveys conducted over a two-year period.

The results of the existing BT accuracy assessments are mixed. The early attempts of conducting BT for cultural goods produced estimates of doubtful accuracy (e.g., transfer errors of a magnitude of several hundred percent in the international BT concerning historic temple sites in Vietnam and Thailand; Tuan et al. 2009) and undermined the method's application in culture. Brown (2004) reported four transfer errors ranging from 4% to 111%, suggesting doubtful usefulness of the BTs associated with the upper end of this range of transfer errors. It is worth noting that Brown (2004) arbitrarily selected four combinations from the whole set of alternatives specified in the experimental design, and the magnitude of transfer errors for all scenarios could have been larger. In turn, some of the later studies (Fujiwara et al. 2018; Lawton et al. 2021b; Lawton et al. 2021c, Wiśniewska et al. 2023), reported very low transfer errors. The majority of the accuracy assessment studies (Mourato et al. 2014; Fujiwara et al. 2018; Lawton et al. 2018) recommended the use of unit value transfer approaches, as they are easy to apply and were found to provide estimates of a comparable or better accuracy than the function transfer approach.

The studies also emphasized that a large similarity of goods between study and policy sites (i.e., homogeneity along characteristics such as significance, physical features, and the range of services provided) enhanced the transferability and, thus, was crucial for the accuracy of BT (e.g., Fujiwara et al. 2018). However, this can be challenging for cultural goods that are often unique and lack perfect substitutes (Provins et al. 2008). Researchers, conducting primary valuation stated preference studies, typically define the evaluated goods using names of specific cultural sites or institutions, which automatically represent integral and unique sets of characteristics difficult to transfer across contexts. To address this issue and mitigate potential threats to the accuracy of BTs, providing respondents in primary stated preference research with descriptions in which cultural goods are described by generic characteristics can be helpful. However, applications following this approach remain scarce. Among the identified BT studies, only Mourato et al. (2014) use estimates from the study in which all of the country's built heritage interiors are conceptualized as a commodity, supporting a similar range of services that are substitutable and homogenous across assets. They find that transfers of such values between two countries (i.e., Germany and the UK) yield smaller TEs than transfers of values elicited for specific heritage assets identified by their names in those countries. In this paper, we adopt a methodology proposed by Mourato et al. (2014) and use estimates for BT derived from a stated preference DCE with generalized attributes to enhance the transferability of our estimates. The estimates we employ can further offer a practical means for valuing cultural policy outcomes

in the context of diverse investment situations, varying in object and size (Wiśniewska and Czajkowski 2017).

We present novel evidence along several dimensions. First, we examine the accuracy of BT when marginal values, as derived from a stated preference DCE, are used for the transfers. This brings a new perspective to the possibility of transferring values of culture, as BT based on marginal values have not yet been used nor investigated for cultural goods. Second, we apply BT to performing arts, which have received limited attention in valuation literature. Third, we expand the small set of locations considered in BT studies in cultural economics by presenting empirical results for Poland. Fourth, we transfer benefits between provinces, whereas the existing studies have only examined transfers across specific sites or countries. Lastly, we demonstrate the potential of transferring values of performing arts, conceptualized as substitutable and homogeneous goods defined by a set of generic characteristics (here, performance types).

2.2 Types of benefit transfer approaches

The common classification of BT approaches involves a distinction between unit value transfers and function transfers (Navrud and Ready 2007; Johnston and Rosenberger 2010). Both types can be further disaggregated into various subcategories. In general, a unit value transfer occurs when an estimated value from a study site is directly applied as an estimate of the value at a policy site. The value can be transferred "as is", leading to an unadjusted transfer, or it can be adjusted in various ways, particularly to account for some differences between the study and policy sites (e.g., a typical adjustment includes controlling for differences in income levels). The function transfer relies on a function defining how the value is affected by various characteristics and next the value at the policy site is calculated based on the relationships described within the function. Functions may be derived from various sources, such as primary valuation data from an individual study or a meta-analysis synthesizing multiple earlier studies. In our empirical application, we consider three types of transfers: an unadjusted unit value transfer; an income-adjusted unit value transfer; and a function transfer.

A common adjustment in unit value transfers includes a correction for mean income differences between study and policy sites. Typically, a constant level of income elasticity of demand is used. Formally, the BT approach can be represented as follows:

$$value_{policy} = value_{study}\left(\frac{income_{policy}}{income_{study}}\right)^{elasticity}. \quad (1)$$

To apply this BT approach, a researcher needs to make an assumption on the level of income elasticity. The literature varies in findings on that elasticity for cultural goods. Seaman (2006) reports the elasticity to range from insignificant and even negative levels to levels substantially above one, as based on a review of econometric studies on demand for performing arts published since 1966. More recent studies do not provide a clearer picture on the elasticity value. Some observe the elasticity to be substantially less than one (e.g., Castiglione and Infante 2016), other report it to be close to one but statistically different from one (e.g., Bernat et al. 2012), and other find it to be approximately equal to one (e.g., Zieba 2009, for disposable income). In turn, findings from other areas, such as health and environmental economics, show that an assumption of income elasticity being equal to one works well in unit value transfers between sites characterized by considerably different income levels (e.g., Barton 2002; Czajkowski and Ščasný 2010; Lindhjem and Navrud 2015). In this study, we follow the latter approach and assume the income elasticity to be equal to one.

For a function transfer, we start with estimating province-specific models, which reveal the relationships how the mean value changes with changes in explanatory variables (such as having a university degree and living in a province capital city). Including the explanatory variables helps us control for differences between the policy and study sites. In the next step, we use the estimates of the coefficients capturing the relationships to predict the value at the policy site. To that end, we insert policy-site population means for the respective explanatory variables to the function describing the value changes.

In our study, the explanatory variables include two (zero-one-coded) binary variables: one controlling for whether a respondent has attained a university degree and another controlling for whether a respondent lives in a capital city of their province. The selection of this functional form is driven by a range of considerations. First, attained education is empirically evidenced to affect preferences towards cultural goods and people's cultural choices (e.g., Seaman 2006; McKenzie and Shin 2020). Second, in Poland, the provinces' capitals are typically the largest local cities, being at the same time local cultural centers. People living in the province capitals typically have the easiest access to theaters and, thereby, may be better familiar with the evaluated good (i.e., theater performances) than those living further away from the capitals. Third, we limit the set of characteristics for econometric purposes, that is, to assure that the model parameters are estimated on a sufficient number of observations. Finally, applications of function transfers require data on policy-site population means of characteristics included in the explanatory variables. Hence, to make function transfers

possible, primary research needs to consider only such explanatory variables for which policy-site population may be available (e.g., provided by public statistical offices).

3 Empirical data and study context

3.1 Discrete choice experiment

The empirical data to the study comes from an online stated preference survey that concerns preferences of Polish residents towards changes in the scope of theater performances offered by theaters in Poland that are subsidized from the state budget. In the season 2017/2018 preceding our survey, there were over 900 theaters in the country, including 121 public venues. On average, a theater provided six performances weekly and two premiers annually. The theaters received approximately 210 million EUR in total public subsidies (coming from local, regional, and central governments), which constituted about 9% of the total expenditures on culture from the public budget. In Poland, all public theaters are subsidized by local and regional governments, excluding four national theaters that exclusively receive funding from the central government budget (Theatre Institute 2019). Although the number of public venues represented only 13% of the total number of theater venues, public theaters clearly dominated the sector by providing 48% of all premiers, giving 94% of all performances, and gathering 97% of all viewers. They also received most of the public subsidies for theaters,[6] which constituted over 80% of their budgets. As of 2018, there were only 387 visitors to performing arts venues per 1,000 residents on average in Poland (Statistics Poland 2023). However, all residents bear the costs of subsidizing public theaters. Value estimates from a DCE study, such as the one described here, can provide input to efficiency assessments of cultural policies and inform if the public expenditures on theaters are justified by social welfare considerations. Specifically, the value estimates can help answer the question of whether the benefits from the theaters, including those to non-theatergoers, outweigh the public expenditures on the theaters.

The main part of the questionnaire, which provides the stated preference data for the analysis, is the DCE. In the DCE, respondents face a sequence of choice tasks and are asked to indicate their preferred option in each task. Every task has two options for choice: one presents a scenario of a possible, hypothetical extension of the scope of theater performances in a respondent's province and the other represents maintaining the current state of the theater performances (i.e., a status quo option).

[6] A small portion of public expenditures on theater is allocated to non-public venues through targeted subsidies.

The scenarios of performance extension are described by four non-monetary attributes (characteristics), which stand for different types of theater performances: entertainment, drama, children's, and experimental. Each of the attributes can take one of three levels: an increase by 50%, an increase by 25%, or no change (all compared to the current state). The current state significantly differs across Polish provinces, and thus the changes consistently expressed in relative terms to all DCE respondents translate into varying absolute changes in the number of performances depending on a province. In addition, the choice options are described by a monetary attribute defined as an individual cost of implementing the performance extension project, which would be collected from everyone in a given province through increased taxes. The possible cost levels range from 5 to 50 Polish zloty (PLN).[7,8] Table 2 summarizes the information about the DCE attributes and their levels.

Table 2. Attributes and attribute levels used in the DCE.

Attribute	**Description**	**Attribute levels**
Entertainment performances	Change in the number of entertainment performances in the province	no change (SQ), 25% more, 50% more
Drama performances	Change in the number of drama performances in the province	
Children's performances	Change in the number of performances for children in the province	
Experimental performances	Change in the number of experimental performances in the province	
Cost	Change in individual expenditures per year (collected through obligatory taxes and fees)	0 (SQ), 5, 10, 20, 50 PLN

Notes: SQ denotes the level used in the status quo (current state) option.

An example of a choice task for one of the provinces is displayed in Figure 1. Each respondent is presented with a sequence of four choice tasks regarding the extension of the scope of theater performances in their province. To mitigate order effects, we randomize the order of the choice tasks and attributes as presented to respondents.[9] The combinations of the attribute levels that define different scenarios of the extension project presented in different

[7] Approximately, 1 EUR is equal to 4.5 PLN.
[8] In the first phase of the data collection (before updating the design; only for 268 respondents), two additional cost levels were used: 2.5 PLN and 25 PLN.
[9] The sequence of DCE tasks with the performance extension includes in fact eight choice tasks, all displayed in a randomized order. Half of them consider the change in a respondent's province, and the other half consider the change in the country as a whole. Given our interest in province-specific value estimates, we only use data from the former in this study.

choice tasks are created in Ngene to optimize D-efficiency of a multinomial logit model, based on data from a pilot study (Ferrini and Scarpa 2007; Scarpa and Rose 2008).

Figure 1. An example choice task for Mazowieckie (MA), translated from Polish.

	Alternative A **Status quo**	**Alternative B**
Theater offer **in province**  **Mazowieckie**		
 Entertainment performances	No change	25%
 Drama performances	No change	50%
 Children's performances	No change	No change
 Experimental performances	No change	50%
Change in your expenditures per year	0 PLN	20 PLN

The questionnaire follows a standard structure for a stated preference survey. It starts with questions needed to meet the intended quotas for gender, age, education level, and province of residency. Next, warm-up questions on the topic of the survey are asked, mostly concerning past visits to theaters and the location (province), where the visits took place. Further, the four types of theater performances are thoroughly introduced. Each type is described on a separate screen with an accompanying symbol to facilitate the performance type understanding and identification (the same symbols are used in the DCE tasks, as shown in Figure 1). The screens provide province-specific examples of performances for each type, which are carefully selected to match theater shows offered in a respondent's province. For each performance type, respondents are also asked whether they have watched a performance of a given type and if yes, they are then asked to provide the performances' titles or locations (theater names). After the explanation of performance types, the DCE is conducted. Initially, respondents are provided with introductory instructions on how the DCE will look like, what attribute levels are possible, and how many choice tasks they will face. To enhance understanding of the attribute levels, the questionnaire provides examples of what a given increase in the scope of the performances would mean in terms of the number of offered premieres and performances. Every choice task is displayed on a separate screen. The final part of the questionnaire includes additional questions about perceptions and opinions, and socio-demographic data.

3.2 Survey administration and development

The survey was administered online by a professional polling agency on the nationwide sample in Poland from May through December in 2018. In total, 2,863 respondents completed the questionnaire, including interviews from a pilot study conducted on 268 respondents. Quotas were used to assure that the distribution of respondents across provinces mirrors the relative sizes of the provinces' populations.

The design of the survey is guided by extensive consultation of the questionnaire with experts from theater, arts, and sociology. Across the literature, there are substantial differences in categorizations of performance types (c.f., Wiśniewska 2019). In our study, we use the categorization into four types as proposed by Wiśniewska and Czajkowski (2017) in their DCE study. The entertainment and drama types are rooted in the ancient division between comedy and tragedy, and they include performances focused on entertainment and classical repertoire, respectively. The children's and experimental types aim to distinguish performances targeted at the youngest audience, often with the educational focus, and experimental forms of theater. Based on information from respondents of the survey of Wiśniewska and Czajkowski (2017) as well as our initial testing of our questionnaire with individuals from the general population, this classification enhances respondents' understanding of the range of performance types offered by theaters and is quick to learn. At the same time, the grouping into four categories may be useful for policymakers in the context of Polish cultural policy.

For a given respondent, the DCE tasks presenting the scenarios of the province-wide performance extension concern only one province. Taking the example of the DCE task from Figure 1, that respondent is displayed with province-wide extension projects only for province Mazowieckie. In other words, we do not provide the same respondent with DCE tasks varying the provinces of implementation of performance extension. We refer to the province used in the DCE as a respondent's province, and it is a province in which a respondent visits theater performances (for theatergoers), intends to do so or is most likely to do so (for non-theatergoers). We identify the respondents' provinces based on early questions in the questionnaire, which query about the frequency and location of their theater visits, both past and planned. For individuals who visit theaters regularly, their province is the one where they attend theater performances most frequently. For those who do not visit theaters but have intentions to do so in the future, their province is the one where they would most likely watch a theater performance. For respondents who neither currently visit, nor have plans to visit theaters, their province is their place of residence. This implies that a respondent could be asked

in the DCE tasks about performance extension in a province different from their province of residence (for example, respondents who regularly attend theater performances outside of their province of residence). For the purpose of this analysis and to assure clarity of the BT investigation, we only include in this study the respondents for whom the province displayed in the DCE tasks is also their province of residence.

3.3 The studied provinces

The empirical analysis is based on data for six provinces of Poland: Lubelskie (LU), Lodzkie (LD), Malopolskie (MA), Mazowieckie (MZ), Wielkopolskie (WP), and Zachodniopomorskie (ZP). The country is divided into 16 provinces, and they represent the highest-level administrative division. We exclude from the analysis the provinces for which we have fewer than 100 observations (to ensure a sufficient number of observations per estimated parameter) and for which econometric models of preferences do not converge credibly (which may be tied to a limited number of observations per province). Figure 2 presents the map of provinces in Poland, where the provinces included in our study are highlighted.

Figure 2. The studied provinces of Poland.

For the six considered provinces, the total number of respondents included in the empirical analysis is 1,328.[10] The distribution of the survey respondents across the provinces, together with basic statistics for the provinces, is reported in Table 3. The number of respondents per province ranges from 111 to 416, reflecting variation in the population sizes of

[10] The number of respondents is 1,167 for function transfers due to 161 missing observations for the indicator variable of living in the province capital city, which is used in the preference models for function transfers.

the provinces. Notably, MZ, with the capital city of the country (Warsaw), has by far the largest population, with more than five million of residents (Statistics Poland 2018). The survey statistics for income and education are also consistent with actual variations along these characteristics between the provinces (Statistics Poland 2019a, 2019b). Based on self-reported data from our survey, the average respondent's monthly net income ranges from 2,394 PLN in LD to 3,405 PLN in MZ. The share of respondents with a university degree spans from 28% in LD to 47% for MZ. Less than half of the respondents in each province reside in the province capitals, except for MZ, where more than 70% declare residing in the capital city. The highest shares of theatergoers, who visit a theater more than once a year, are found in MZ and MA, where they reach the levels of 47% and 46%, respectively. These percentages correspond to the sizes of the theater markets of the provinces. The country's capital city is located in MZ and boasts the largest theater market in the country. The second largest theater city in Poland is Cracow located in MA. We also note that the study sample includes substantial shares of respondents who never go to theaters or attend theater performances rarer than once a year (36% of the sample) and of those who do not plan to go to theater in the future (20%), which helps us better capture the variation of preferences in the society.

Table 3. Survey-based statistics for the studied provinces.

Province	Number of respondents	Average monthly net individual income [PLN]	Respondents with a university degree	Respondents living in the capital of their province	Respondents that visit theaters more than once a year
Lubelskie (LU)	144	2,412	42%	42%	34%
Łódzkie (LD)	197	2,394	28%	47%	40%
Małopolskie (MA)	238	3,292	42%	42%	46%
Mazowieckie (MZ)	416	3,405	47%	74%	47%
Wielkopolskie (WP)	222	2,557	37%	38%	37%
Zachodniopomorskie (ZP)	111	2,299	40%	27%	41%

Notes: In total (for the six provinces jointly), there are 109 missing observations for income and 161 missing observations for the indicator of living in the province capital city.

The studied provinces are distributed between the east (LU, LD, MA, MZ) and west (WP, ZP) of the country, with a greater concentration in the east. The so-called Polish eastern wall has been recognized as a peripheral region with a weaker economic position compared to the rest of Poland (Bartosiewicz and Stańczyk 2014). One of the reasons for this lies in the region's agricultural character. However, the division into the east and west has persistent historic grounds: the specificity of the 17th-19th-century partitions led to a lower level of modernization

in the east of Poland, resulting in underdeveloped infrastructure, a lower degree of industrialization and urbanization, a strong prevalence of traditional family farming, and a relative neglect of education and science in provincial budgets (Kolasa-Nowak 2011). The recent data on cultural participation mirrors these differences across regions. Specifically, despite MA and MZ being the most important theatrical centers in the country, official statistics show that cultural participation in the east of Poland (including participation in performing arts) is lower than in the west of Poland (Statistics Poland 2020).

In Appendix Table A1, we present additional information for the comparison of theater markets across the provinces, focusing on data from the year 2018, sourced from Statistics Poland (2023). Local public expenditures on theaters per capita in that year varied from 23.70 PLN in WP to 38.73 PLN in MZ which outperformed MA as the province with the second-largest expenditure by more than 10 PLN per capita per year. Besides the two largest theater centers, MA and MZ, the number of public theater venues varied from five in LU to 11 in WP. In MZ and MA, 27 and 20 public venues were located, respectively. It is evident that MZ, with the country's capital city, significantly outperformed other provinces in terms of both public expenditures on theaters and the number of theater venues, among others. The disparities between provinces are even more pronounced when looking at the number of performances shown in 2018. It varied from 704 in LU to over 10,000 in MZ, with a median of 3,599 for all six provinces under study. Again, it is worth noting that MA, as the second-largest province in terms of the number of performances, had over 5,500 performances fewer than MZ.

One of the factors that determine the effectiveness of cultural services provided by public institutions, such as public theaters, is their accessibility to the public. Research shows that theater attendance depends on the number of available seats (Smolny 2017). When accounting for the differences in population size between provinces, interesting insights are provided by the number of residents per one seat in performing arts institutions, indicating higher accessibility of theater services per resident in provinces like MZ with 269 and ZP with 382 residents per seat, compared to WP with 610 residents per seat, and LU with more than 1,000 residents per one seat, suggesting the lowest accessibility across provinces. It should be highlighted that public theaters are primarily located in the largest cities and metropolitan areas, which are usually the provinces' capitals. Therefore, this measure of theater accessibility does not fully illustrate the spatial distribution of residents' access to theater venues within a province.

4 Empirical strategy

We start with estimating marginal values based on the primary data, which is described in the earlier section. This provides us with primary value estimates that are subsequently used for BTs and for the assessment of the BTs' accuracy.

We follow a standard approach to DCE data modeling, as used in numerous applications of the methodology (e.g., Browning et al. 2024; Czajkowski et al. 2022). Similarly, we apply standard procedures for BTs and the assessment of their accuracy (e.g., Johnston and Zawojska 2020). This way, our empirical assessment is performed with the use of typical approaches in this area.

4.1 Econometric models

Modeling of preferences stated by respondents in DCE tasks is grounded in a random utility framework (McFadden 1974). According to the framework, the utility of individual *i* from policy scenario *p*, $U_{ip}(\cdot)$, is a function with two main additive components: observed characteristics of the policy, which are described through the non-monetary and monetary DCE attributes (that is, $\boldsymbol{X}_{ip}$ and C_{ip}, respectively), and unobserved idiosyncrasies expressed as an unobservable error term, ε_{ip}. Formally, the utility function can be represented as:

$$U_{ip}(\cdot) = \boldsymbol{\beta}_i' \boldsymbol{X}_{ip} + \alpha_i C_{ip} + \varepsilon_{ip}, \tag{2}$$

where $\boldsymbol{\beta}_i$ and α_i are preference parameters to be estimated, which represent marginal utilities from the non-monetary and monetary attributes, respectively. Indexing the parameters over *i* denotes that the parameters are allowed to vary over individuals according to a predefined multivariate distribution (Train 2009). This accounts for heterogeneity of preferences across the individuals and produces a random-parameter specification of the model, which is typically (and henceforth) referred to as a mixed logit model (MXL).

We estimate the model in willingness-to-pay (WTP) space to ease the interpretation of the model results. In a WTP-space model, the preference parameters can be readily interpreted as monetary WTP amounts. To that end, we first define $\alpha_i = \gamma_i / \theta_i$ as the monetary parameter in preference space, where γ_i represents the underlying marginal utility of income and θ_i is a scale parameter, and we define $\boldsymbol{\beta}_i = \boldsymbol{\lambda}_i / \theta_i$ as a vector of preference-space parameters on the non-monetary attributes, where $\boldsymbol{\lambda}_i$ denotes a vector of underlying marginal utilities associated with these attributes. We further assume error term ε_{ip} to follow an i.i.d. type I extreme value distribution with a constant variance of $\pi^2 / 6$ (Scarpa et al. 2008; Train and Weeks 2005). A

marginal WTP value (i.e., an implicit price; $\boldsymbol{\omega}_i$) for a change in a given non-monetary attribute can be calculated as a ratio of the attribute coefficient to the coefficient of the monetary attribute—that is, $\boldsymbol{\omega}_i = \boldsymbol{\beta}_i / \alpha_i = \boldsymbol{\lambda}_i / \gamma_i$. Based on that we can re-define preference-space model specification (2) into a behaviorally equivalent specification, but in WTP space:

$$U_{ip}(\cdot) = \alpha_i \left(\frac{\boldsymbol{\beta}_i}{\alpha_i}' \boldsymbol{X}_{ip} + C_{ip} \right) + \varepsilon_{ip} = \alpha_i \left(\boldsymbol{\omega}_i' \boldsymbol{X}_{ip} + C_{ip} \right) + \varepsilon_{ip}. \quad (3)$$

The WTP parameters in $\boldsymbol{\omega}_i$ are allowed to vary across individuals according to a predefined distribution.

In our empirical application, we assume the elements in $\boldsymbol{\omega}_i$ to be normally distributed. We define $\alpha_i = e^{\upsilon_i}$, with υ_i being an underlying latent normal factor that defines the lognormal distribution of the coefficient of the monetary attribute. This assumption follows from a common practice in discrete choice models to ensure a positive marginal utility of income. In our application, vector $\boldsymbol{X}_{ip}$ also includes a constant for the status quo, equal to one for the status quo option and zero otherwise, in order to measure the marginal utility and the associated WTP for maintaining the current state of theater performances. We refer to this model specification as a basic MXL.

In addition, for the function transfer approach, we extend the basic model specification in (3) to include variables helping capture the effects of selected demographic characteristics on the marginal WTP values. In this extended model specification, the means of the random WTP parameters are interacted with the demographic variables. The estimates on the interaction terms express by how much the marginal WTP values differ for individuals characterized by specific demographic factors. We call this model specification an extended MXL.

4.2 Benefit transfer procedure

The estimates from the basic models are used for unit value transfers, and the estimates from the extended models are used for function transfers. We consider all possible combinations of pairing the provinces into the sets of a study site and a policy site. This means that in separate BTs, each province is treated as a policy site and the value estimates derived from all other provinces are transferred to that policy site–these other provinces are then considered as study sites.

In the unadjusted unit value transfer, we take the WTP estimates as obtained from a basic MXL model for a given province and transfer them to a province assumed to be a policy site. In the adjusted unit value transfer, the WTP estimates derived from the basic MXL models are scaled by the ratio of mean income levels of the provinces between which the transfer is conducted (as per equation (1)). In the function transfer, we transfer the estimated function to the policy site and calculate the transferred value based on it. For example, let us assume that the marginal WTP values for extending drama performances are described by the following set of estimates: the mean parameter estimate of extending drama performances is equal to β_1^D, the parameter estimate of the interaction term of this performance type and the variable controlling for living in the province capital is equal to β_2^D, and the parameter estimate of the interaction term of this performance type and the variable controlling for having a university degree is β_3^D. Then, the transferred value to a policy site will be the sum of β_1^D, β_2^D multiplied by the share of people living in the province capital at the policy site, and β_3^D multiplied by the share of people having a university degree at the policy site.

4.3 Accuracy assessment of benefit transfers

Examination of the accuracy of BT relies on a comparison of the true, actual value at a policy site and the transferred value to the policy site, where the latter is derived from the value estimate(s) based on primary data for other site(s) (i.e., study site(s)) as described in the preceding paragraph. In a standard BT application setting, one does not know the actual value at the policy site and, hence, the BT is needed. In our study, we treat as the true, actual values the WTP values estimated directly from the DCE data for a given province. In other words, we treat these primary-data-based values as criterion values for the accuracy assessment (a similar approach is employed in the existing BT accuracy assessment studies reported in Table 1; i.e., Tuan et al. 2009, Mourato et al. 2014, Fujiwara et al. 2018, Lawton et al. 2021b, Lawton et al. 2021c). We separately assess two components of accuracy–validity and reliability–as per framework by Bishop and Boyle (2019).

The validity assessment requires the evaluation whether there is no statistical difference between the actual value and the transferred value. To that end, we conduct Z-tests with the null hypothesis of equality between the actual and transferred marginal values. A valid transfer is when the two values are statistically indistinguishable (i.e., no reason to reject the null hypothesis).

The reliability assessment employs a commonly used measure of transfer error (TE; Rosenberger and Stanley 2006). It helps understand how close the transferred value is to the actual value. We use the absolute value percent TE:

$$|TE| = \frac{|value_{transferred} - value_{actual}|}{value_{actual}} \cdot 100\%. \quad (4)$$

The lower the TE, the more reliable the transfer. Henceforth, for brevity, when we use the term TE, we mean the absolute value percent TE as defined in equation (4).

We argue that both components of the BT performance assessment–validity and reliability–are needed and complementary (and have been used in earlier examinations, e.g., Johnston and Zawojska 2020). In particular, the probability of failing to reject the null hypothesis of equality can be larger for WTP with large standard errors than for those with small standard errors. This implies that WTP estimates with large standard errors can be expected to be more likely to generate valid transfers. Then, the assessment of reliability can be helpful to identify accurate transfers. As also observed by Rosenberger (2015) based on reviewed BT inquiries, the correlation between passing validity tests and minimizing TE is low, which further points to the importance of examining both validity and reliability. We further note that the two angles of the accuracy assessment have a very different character (Rosenberger 2015). The accuracy assessment provides a strict condition, defined by the statistical test, on whether a BT is valid or not. On the other hand, the reliability assessment does not set such a definite line between reliable and unreliable transfers, but rather provides a general indication about a BT performance.

5 Results and discussion

5.1. Primary value estimates and illustration of the benefit transfer procedure

Primary value estimates to our BTs come from two sets of models: (i) basic MXL models only with the variables controlling for the choice option characteristics used in the DCE and (ii) extended MXL models including in addition interactions of the mean coefficients with the variables controlling for selected demographic characteristics (i.e., having a university degree and living in a province capital). For each province, we estimate a separate basic model and a separate extended model, using the simulated maximum likelihood method with 4,000 Sobol draws.[11] The estimates from the basic models are used for unit value transfers (both unadjusted and income-adjusted), and the estimates from the extended models are used for function transfers.

[11] The estimation is conducted based on a custom code developed in Matlab available at https://github.com/czaj/DCE under CC BY 4.0 license.

Table 4 presents the main results of the basic MXL models. For brevity, in the table, we report only the estimates of the coefficients on the means, and full results of the models are presented in Appendix Table A2. Since the models are estimated in WTP space, the non-cost coefficient estimates can be directly interpreted as WTP values in PLN per person per year. Specifically, the coefficients for the non-monetary DCE attributes represent mean marginal WTP values for a one-percent increase of the number of performances of a given type. The estimates on the monetary attribute do not have any direct interpretation.

Table 4. Mean value estimates from the basic MXL models in WTP space (for unit value transfers).

	Model for LU		**Model for LD**		**Model for MA**		**Model for MZ**		**Model for WP**		**Model for ZP**	
	Mean effects coefficients (standard errors)											
Status quo	0.33	***	0.24	***	0.29	***	0.19	***	0.15		0.12	***
	(0.04)		(0.09)		(0.01)		(0.00)		(0.12)		(0.03)	
Entertainment	-0.04		0.35	***	0.19	***	0.09	***	0.42	*	0.05	
	(0.06)		(0.13)		(0.05)		(0.01)		(0.23)		(0.05)	
Drama	0.16	***	0.17		0.06		0.19	***	-0.25		-0.08	
	(0.05)		(0.14)		(0.05)		(0.01)		(0.23)		(0.08)	
Children's	0.21	***	-0.02		-0.37	***	0.26	***	-0.09		-0.11	**
	(0.06)		(0.12)		(0.06)		(0.01)		(0.21)		(0.05)	
Experimental	0.12		0.10		0.07		-0.11	***	-0.38	*	-0.14	**
	(0.09)		(0.14)		(0.05)		(0.01)		(0.21)		(0.07)	
Cost (100 PLN)	-3.55	***	-2.99	***	-3.91	**	-3.40	**	-2.10	***	-4.41	***
	(1.13)		(0.96)		(1.53)		(1.76)		(0.58)		(1.66)	
Log-likelihood at convergence	-290.54		-343.86		-458.09		-820.33		-446.15		-185.76	
Number of observations (n)	576		788		952		1664		888		444	
Number of respondents	144		197		238		416		222		111	

Notes: ***, **, and * represent significance at the 1%, 5%, and 10% level, respectively. The variables are modeled as random and normally distributed, except for the cost, which is assumed to follow a lognormal distribution and the estimates of the underlying, preference-space equivalent, normal distribution are provided. To ensure the model convergence, the cost is divided by 100. Full results of the models, including standard deviations and model characteristics, are reported in Appendix Table A2.

In all, but one (WP), provinces, the variable on the status quo is statistically significant and positive, indicating that people have some WTP to maintain the current state of the scope of theater performances offered in their provinces. At the same time, they also like some of the considered extensions of the number of theater performances. In particular, respondents are willing to pay to increase the number of entertainment performances in all, but two (LU and ZP), provinces, and the marginal WTP values for a one-percent increase in that number range from 0.09 to 0.42 PLN per person per year. In LU and MZ, they are willing to pay about 0.16-0.19 PLN to have the number of drama performances increased by one percent. The extension of children's performances is associated with mixed preferences across the provinces. While respondents in LU and MZ would like to pay 0.21-0.26 PLN for a one-percent increase of these performances, respondents in MA and ZP would need to be compensated 0.37 and 0.11 PLN, respectively, for a similar change. In turn, in none of the considered provinces, respondents have a positive WTP for extending the number of experimental performances. Respondents in MZ, WP, and ZP would need to be paid a compensation of 0.14-0.38 PLN for a one-percent increase of the number of this type of performances.

Interestingly, the obtained WTP values (when statistically significant) are consistently of the same sign across all six provinces for some types of performances (specifically, positive WTP values for entertainment and drama performances, and negative WTP values for experimental performances), while they have varying signs for the other type (that is, children's performances). We hypothesize that these differences arise as a result of the specific nature of the latter performances, as the population as a whole has less familiarity and experience, potentially contingent on the fact of being a parent or planning to have children in the future. Many performances for children are tailored for schools. Consequently, people attend them as pupils, and the attendance in the early ages does not necessarily translate into familiarity with this type of performances among adults, from which our respondents are recruited.

he extended MXL models with the demographic variables illustrate how the marginal WTP values change for different groups of individuals–that is, those with and without a university degree and those living in and outside of their province capital. These results are reported in Appendix Table A3. The mean value estimates remain fairly robust across the basic and extended models. In the extended model, these mean estimates represent the WTP values for respondents without a university degree and not living in a province capital. The two demographic characteristics affect WTP to different degrees and directions for various

provinces. This suggests a lack of systematic effects of these characteristics on WTP for the considered extensions of theater performances. For example, for all provinces except for MA, we find that respondents living in the provinces capitals are characterized by a lower WTP for extending the number of entertainment performances than those living outside of them, and an opposite effect is observed for MA (that is, a higher WTP among those living in the province capital). Those living in the provinces' capitals are willing to pay more for more drama performances in LU, LD, and WP, but they are willing to pay less for more drama performances in MA and ZP. Differences such as these evidence province-specific nature of the effects.

The results described above provide primary value estimates to our study of BT accuracy. For illustration of how we conduct the BTs, we take the example of transferring marginal WTP values for increasing the number of entertainment performances (by one percent) from MZ to MA. In the unadjusted unit value transfer, we simply use the estimate for MZ, which is equal to 0.09 PLN as reported in Table 4, and assume that this is the transferred value in MA. The actual value in MA is 0.18 PLN, and we use it to assess the transfer accuracy. Specifically, based on a Z-test, the two values are statistically different and so we conclude that the transfer is not valid. We measure the reliability with the TE, which is equal in this case to |0.09-0.19|/0.19 = 53%. In the income-adjusted unit value transfer, the transferred value is adjusted by the ratio of the mean income levels in the two provinces. Specifically, we use equation 1 and calculate that the transferred value is equal to 0.09*(3,292/3,405) = 0.09 PLN, where the income levels are taken from the statistics reported in Table 3. Finally, to conduct the function transfer, we use the marginal value estimates for MZ from Appendix Table A3 (for both the main effect and two interaction terms) and the mean statistics for MA for the demographic variables as reported in Table 3. Specifically, the calculation of the transferred value is 0.33-0.28*0.42+0.04*0.42 = 0.23 PLN. For each of the latter cases (the income-adjusted unit value and function transfers), we do similar assessments of accuracy as illustrated for the unadjusted unit value transfer.

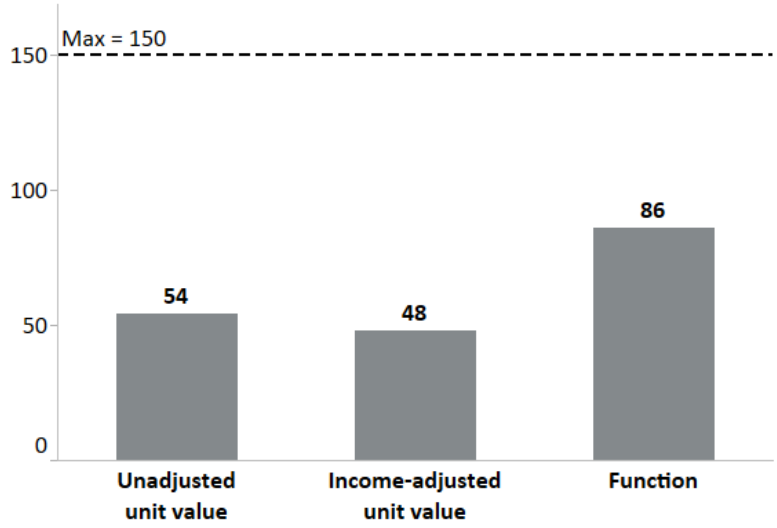
5.2. Overall accuracy of benefit transfer

In this section, we describe the overall results for validity and reliability of the BTs we conduct. We perform a total of 150 transfers of marginal values for each of the three considered BT approaches. Validity is assessed based on Z-tests, where the null hypothesis assumes that the transferred value estimate is equal to the actual value estimate directly derived from the data for a given (policy) site. The hypothesis is evaluated based on a 5% significance level. Failing

to reject the null hypothesis suggests that the compared estimates are statistically identical and, thus, the transfer can be considered valid.

The results of the validity assessments are presented in Figure 3. We find that the numbers of invalid transfers, yielding statistically different transferred and actual WTP values, appear to be much lower for both types of the unit value transfers than for the benefit function transfer. In particular, we observe 54 (out of the 150) invalid transfers for the unadjusted unit value transfer (which means 96 valid transfers), 48 for the income-adjusted unit value transfer, and 86 for the function transfer. The shares of invalid BTs range from 32% (for the income-adjusted unit value transfer) through 36% (for the unadjusted unit value transfer) to 57% (for the benefit function transfer).

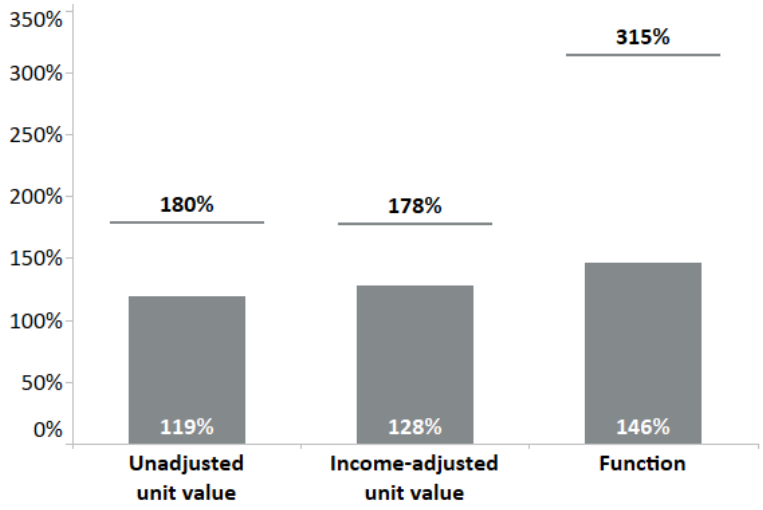
Figure 3. Numbers of invalid BTs.



Evidence from the BT literature suggests that validity in value estimates is rejected most of the time (Rosenberger 2015). However, in our case, the rejection rates of validity for both unit value transfers are lower than 40%, indicating that transferring marginal WTP values of theater performances can be considered valid. In turn, more than half of validity tests are rejected when the function transfer is used.

In Figure 4, we present the means and medians of the TEs. We observe that unit value transfers outperform the function transfer approach in terms of reliability. According to Bateman et al. (2011), this is likely to happen when the commodities in question are highly similar. Since the definition of goods does not vary between provinces in our study, we may indeed consider this as a plausible explanation for our observation.

Figure 4. Mean and median TEs.



Notes: The TE means are represented by the horizontal lines. The TE medians are represented by the bars with the specific number reported on the bottom.

In our study, for the unadjusted unit value transfer, TEs range from low (8%) to very high (>13,000%), with a mean of 180% and a median of 119%. For the income-adjusted unit value transfer, TEs also vary from negligible (5%) to very high (>10,000%), with a mean of 178% and a median of 128%. Regarding the function transfer, we find a larger variation with TEs ranging from no error (0%) to very high (>23,000%), with a mean of 315% and a significantly lower median of 146%. It is important to note that these are transfers of single coefficients that produce outlier TEs and significantly influence the aggregated mean. Among the function transfers, we identify 25 transfers yielding very high TEs, which exceed 500%. For both unit value transfer approaches, we only observe 10 such observations.

In most provinces, the high mean TEs are primarily driven by transfers of single coefficients recognized as outliers. This is illustrated by the presence of long upper whiskers in the boxplots presented in Appendix Figure A1, especially in the case of function transfers. For example, when transferring the value associated with an increase in drama performances using the unit value approach from MZ to LD, we observe a TE of 9%. However, when the value associated with an increase in children's performances is transferred between these provinces, we observe an excessively higher TE of 1133%. Not surprisingly, the high mean TE for this pair is largely influenced by the latter.

While detecting and dropping outlier observations would naturally result in significant reductions in TEs' magnitudes, thereby making the distribution of TEs less skewed, including outliers in the analysis could aid in identifying BT procedures most likely to decrease TEs,

motivating us not to drop outliers. However, focusing on the mean TEs for provinces in which even one outlier is identified risks concerns about the reliability of all transfers within those pairs of provinces, even though transfers of the remaining coefficients produce relatively low TEs. Therefore, the further analysis is based on the median TEs for each province. Medians have been used in the related literature (e.g., Kaul et al. 2013, Johnston et al. 2015a) as they are less sensitive to the presence of outlier observations.

According to Rosenberger (2015), statistical validity is not always required for transfers to be reliable, as it is possible, in some cases, that TEs may fall within an acceptable range despite a failure to achieve statistical validity. Based on our data, we observe that the rejection rate of estimates' equality is positively and statistically significantly correlated with the median TEs when all three BT approaches are considered jointly, however, the correlation is not strong ($r=0.11$, p-value<0.01). When the BT approaches are considered separately, we find no statistical correlation for unit value transfer approaches ($r<0.1$, p-value>0.1), but a statistically significant correlation for the function approach ($r=0.13$, p-value<0.01).

5.3. Accuracy of benefit transfer by DCE characteristics

In the preceding section, we consider the accuracy of transfers in aggregate. This procedure can be handy when evaluating the public theater sector as a whole. However, it is often the case that policies target institutions depending on the type of performances they specialize in (e.g., providing additional funding for theaters focused on experimental performances, since revenues from ticket sales constitute a negligible portion of their budget). In such situations, for example, considering the provision of theater performances by type can support cultural policy makers with greater precision. In this section, we analyze whether the numbers of valid and invalid transfers, as well as the magnitude of TEs, significantly differ across performance types considered in our study.

Figure 5 illustrates that transfers of values associated with the extension of drama theaters' offer can be considered the most valid when unit value transfers are employed. For unadjusted unit value transfer, we observe that the number of invalid transfers (statistically different value estimates) ranges from six (drama performances) to 14 (children's performances). Income adjustment appears to reduce variation in validity assessment across types. We observe that using the income-adjustment unit value approach increases the validity of transfers associated with children's and entertainment performances. The support for the extension of performances is likely influenced by both current and anticipated consumption levels. Participation costs vary across performance types. Entertainment performances

generally feature higher ticket prices and are presumed to have a higher price elasticity of demand. Attending children's performances, in turn, requires parents to allocate extra time and financial resources, as they may need to make time to visit the theater in the busy parental schedule or purchase tickets for both their child and themselves. Therefore, the demand for children's and entertainment performances may be more sensitive to changes in income compared to drama and experimental performances. Consequently, the validity of transfers involving these two types can be more reactive to income-adjustments. The function transfer approach results in decreased validity for each DCE characteristic in our case, with the most significant decline observed in validity when transferring the value associated with drama performances. However, there is also a noticeable deterioration in the validity of transfers involving children's performances and the value of maintaining the status quo.

Figure 5. Numbers of invalid BTs by DCE characteristics.

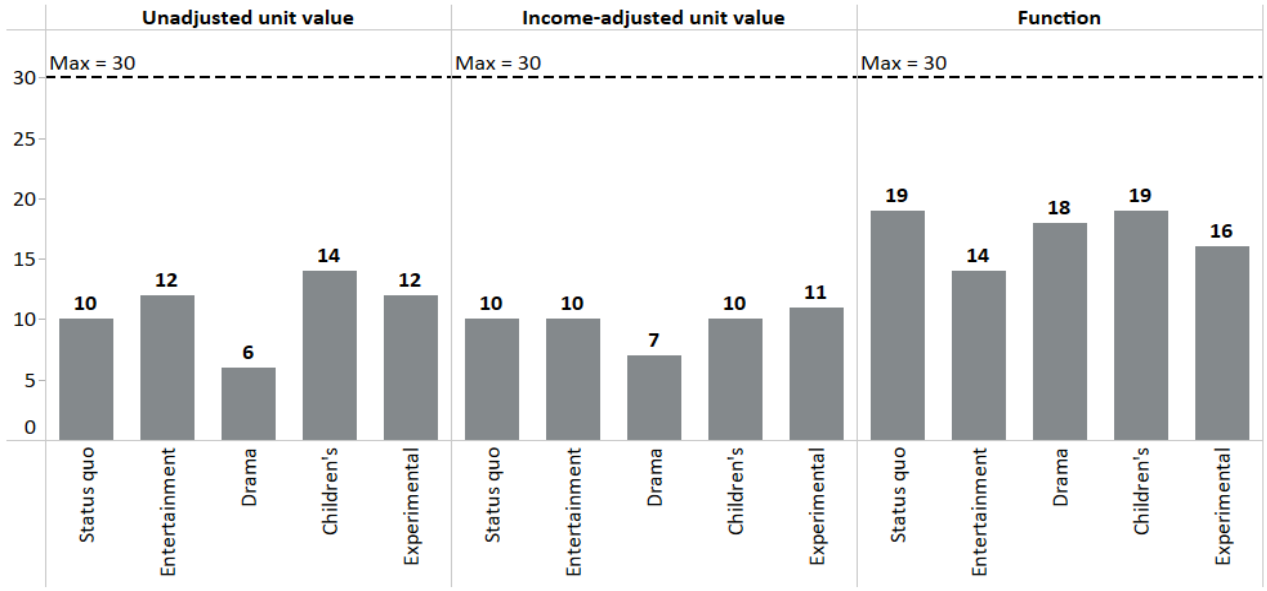
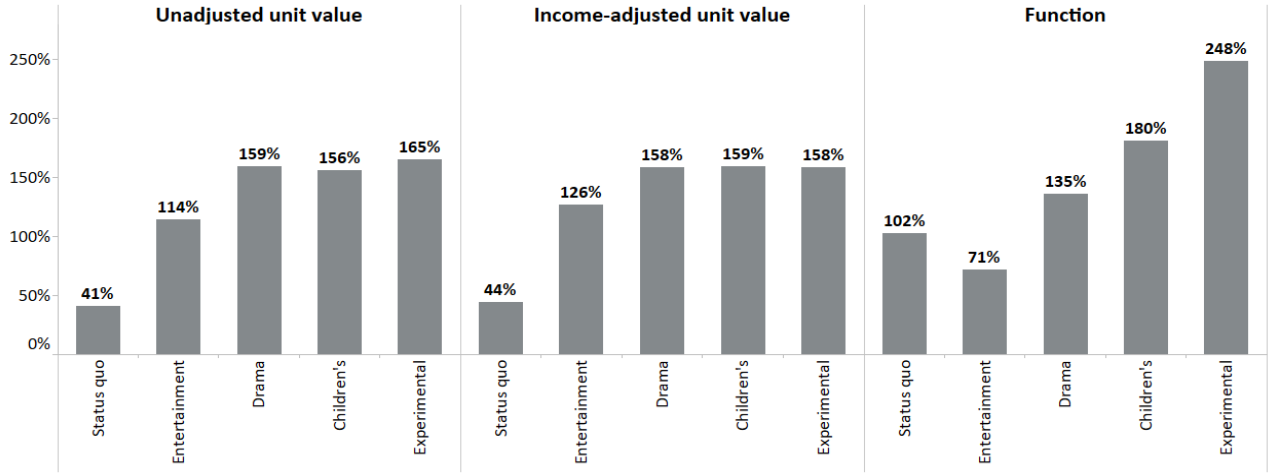


Figure 6 displays the median TEs, where lower TEs for a given DCE characteristic suggest more reliable transfers involving this particular characteristic. The most noticeable aspect when using value approaches is that transfers of the value respondents hold for the current status quo are associated with significantly lower TEs (more than twice as low) compared to when specific policy characteristics are considered, i.e., extensions in a particular type of theater performance.

Figure 6. Median TEs by DCE characteristics.



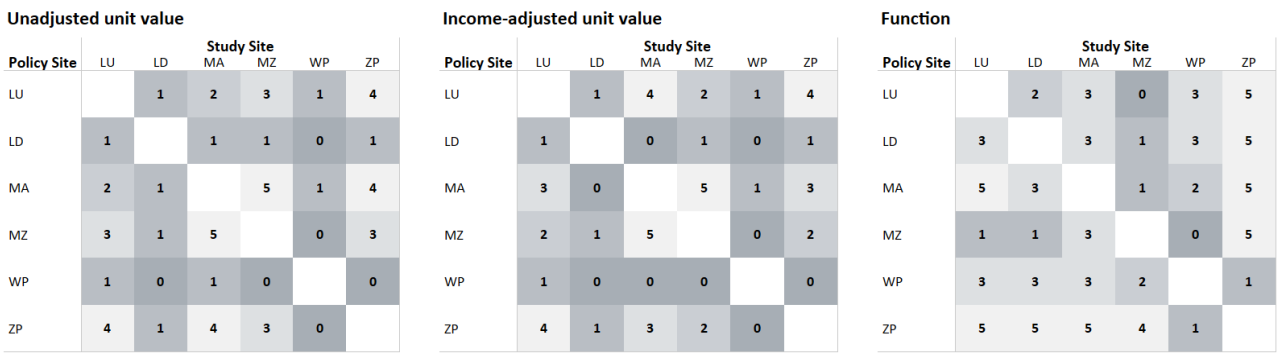
Among the four types of theater performances analyzed, transfers related to an increase in the offer of entertainment theaters seem to be the most reliable, regardless of the BT approach chosen. While no significant differences can be observed due to income adjustments, the function transfer approach leads to significantly increased TEs for children and experimental performances. However, for other performance types, namely drama and entertainment, using this approach results in improved reliability. Particularly, we observe that the inclusion of sociodemographic covariates in modeling decreases TEs for transfers involving entertainment performances, and to a lesser extent, drama performances. While theaters providing drama and entertainment performances are more well-known and commonly experienced by the general public than other theaters (see Wiśniewska and Czajkowski 2019), our functional adjustment accounts for the factors of high education and residing in a large city (province capital), which are typical characteristics of frequent theater goers. Hence, it is possible that the functional form used here decreases TE magnitudes for these performance types by identifying their users. This finding suggests that the accuracy of BT in performing arts may significantly depend on the performance type, with the more well-known types, involving entertainment and drama performances, being associated with higher reliability.

5.4. Accuracy of benefit transfer by provinces

The aggregated results suggest that the transfers of values associated with publicly provided theater services might be considered valid and reliable, and thus be used in policy analysis. However, given substantial heterogeneity across provinces with regard to the size and effectiveness of theater services currently provided by the public sector (as described in Section 3.2), we further analyze the accuracy of transfers for each pair of provinces separately.

The graphs in Figure 7 display the magnitude in the number of tests that were rejected at the 95% confidence level for each pair of provinces. Value estimates are transferred from a province in a column (study site) to a province in a row (policy site). The lower the value, the more valid the transfer, with a value of 0 meaning that all estimates are statistically indistinguishable, suggesting that the transfer from one province to another is fully valid based on individual tests of model coefficients, i.e., for all five DCE characteristics, including maintaining the status quo. Values ranging from zero to three are gradiently marked and indicate valid transfers between particular provinces. Values from four to five, on the contrary, indicate that more than half of the individual model coefficients are statistically different, and thus a given transfer should be considered invalid. For example, for an unadjusted unit value approach, we reject three out of five validity tests when transferring from ZP to MZ. As validity tests only verify the equivalence of two welfare measures, their results for the unit value approach do not depend on the transfer direction. It means that we can observe the same result of three out of five rejected validity tests when transferring from MZ to ZP. Thus, the values are always symmetrically distributed with respect to the diagonal for the unit value approach. The income-adjustment slightly influences the validity assessment for six provinces, but it only influences a single test in each of them. Since we do not observe a significant improvement in any pair of provinces, we can state that the validity of our estimates is not sensitive to differences in income across provinces. Function transfers use a function derived from a primary study to calculate value estimates calibrated to selected characteristics (in our case higher education and place of residence) of a policy site. Thus, the results for the function approach will more likely depend on their direction. For example, we can observe only one out of five rejected validity tests when transferring from MZ to MA, but three out of five rejected when transferring the other way around: from MA to MZ.

Figure 7. Numbers of invalid BTs by provinces.



We observe the highest level of equality in study site estimates and policy site estimates in all unadjusted unit value transfers involving WP (with only 8% of equality tests rejected). Following that, LD exhibits a rejection rate of 16%, whether serving or receiving estimates. While these two provinces demonstrate the best validity results, they also represent, in many ways, the average characteristics within the sample (see Appendix Table A1). When the sites involved in a transfer are averaged in terms of relevant characteristics, the transferred value estimates are more likely to be statistically identical across the study and policy context (Rolfe et al. 2015). On the other hand, the smallest number of valid unit value transfers is observed for MA, with slightly over half (52%) of equality tests being rejected. Among the three remaining provinces (LU, MZ, ZP), the null hypothesis rejection rates are lower than 50%, ranging from 44% to 48%.

Income adjustments do not significantly alter the results regarding validity assessment (see Figure 7). When comparing the benefit function-based estimates to both unit value transfer approaches, the equality of study and policy site estimates show worsened statistical validity in most cases. Rejection rates for function transfers, with respect to the results for income-adjusted unit value transfers, are higher by 46 percentage points for LD, 42 for ZP, 34 for WP, 18 for MA, and 16 for LU.

For four provinces (LU, MA, ZP, LD), the use of the function approach results in more than half of the equality tests being rejected, with ZP being an extreme example, with 82% of transfers failing to achieve statistical validity. Despite the function transfer approach generally leading to a dramatic decrease in validity assessment for five out of six analyzed provinces, we observe its better performance in the case of transfers involving MZ. MZ stands out significantly among the other provinces in terms of two socio-demographic variables used in the extended models (see Table 3). Although not systematically across provinces, both variables influence the way people respond to the theater offer. Education is undoubtedly one of the main determinants of theater attendance and is crucial for the learning-by-consuming process that enhances the perceived value of theater experiences. The second variable indicates living in the province's capital city which, in the case of MA, is Warsaw, the capital city of the country. Living in Warsaw certainly makes people much more exposed to theater performances and, as a result, enhances their familiarity with these relatively niche cultural goods. The improved reliability of the function transfer for MZ suggests that demographic adjustments can perform well (i.e., improving reliability) for a province in which differences with respect to these characteristics are relatively large.

Figure 8 displays a comparison of median TE magnitudes for all six provinces. Again, estimates are transferred from a province in a column (study site) to a province in a row (policy site). Unlike validity tests that verify the equivalence of two welfare measures, reliability tests highly hinge on the direction of the transfer regardless of the approach used. Thus, we observe, for example, a median TE of 63% when performing unadjusted unit value transfers from ZP to WP, but 171% when transferring the other way around: from WP to ZP. Again, the lower the value, the more accurate the transfer can be considered in terms of its reliability, with a TE of 0% meaning that the mean WTP value in the assumed study site (i.e., the true value) is equal to the mean WTP value for the assumed policy site. Based on a review of empirical findings from the BT literature (Rosenberger 2015), in which TEs ranged from very low (<1%) to very high (>7000%) with a mean of 140%, we assume the acceptable range of TEs in our study from 0% to 140%, and thus transfers between particular provinces with median TEs lower than 140% are marked.

Figure 8. Median TEs by provinces.

Unadjusted unit value

Policy Site \ Study Site	LU	LD	MA	MZ	WP	ZP
LU		26	61	43	260	155
LD	35		45	75	244	144
MA	121	83		169	122	74
MZ	74	110	116		232	43
WP	131	71	118	79		63
ZP	183	169	179	76	171	

Income-adjusted unit value

Policy Site \ Study Site	LU	LD	MA	MZ	WP	ZP
LU		25	71	59	251	156
LD	34		60	82	224	144
MA	135	97		167	186	64
MZ	146	114	124		276	77
WP	134	75	114	84		61
ZP	183	169	136	43	155	

Function

Policy Site \ Study Site	LU	LD	MA	MZ	WP	ZP
LU		39	110	33	146	177
LD	35		38	23	178	184
MA	52	64		88	106	162
MZ	43	68	90		84	207
WP	187	197	188	225		154
ZP	243	252	283	331	119	

For the unadjusted unit value approach, we observe 20 transfers (from one province to another, without decomposing into specific attributes) with a median TE lower than 140% and 10 with a median TE higher than that. The lowest median TE (<50%) is found when transferring from LD to LU (26%), LU to LD (35%), ZP to MZ (43%), MZ to LU (43%), and MA to LD (45%). In general, we find LU, LD, MA, and MZ providing the most reliable transfer estimates for other provinces. Four out of five transfers when each of these provinces is assumed as the study site generate median TEs lower than the threshold of 140%. The best province in terms of receiving value estimates from other provinces is WP. All five transfers with this province as an assumed policy site generate median TEs lower than 140%. However, at the same time,

most transfers from WP to other provinces fall beyond an acceptable range of TE, suggesting that the province performs very well when receiving value estimates, and very bad when serving estimates. Income adjustments only slightly change the results and do not help to find any clear pattern in the geographical distribution of median TEs.

For the function transfer, the cluster, including LU, LD, MA, and MZ, crystallizes. These are the provinces that can be considered belonging to the eastern part of Poland. Half of these transfers involving these provinces yield median TEs below 50%, and only one out of 12 has a median higher than 100% (from MA to LU). While functional adjustment (with regard to higher education and residence) do not significantly change median TEs for LU and LD as assumed policy sites (these are also low when using unit value approaches), the functional adjustments significantly correct transfers from these two to MA and MZ, as well as transfers between MA and MZ themselves. Median TE for transfers to MA (assumed policy site) drops by 79 percentage points when transferring from MZ, 33 when from LD, and 83 when from LU, all compared to income-adjusted unit value transfers. Median TE for function transfers to MZ drops by 34 percentage points when transferring from MA, 46 when from LD, and 103 when from LU. At the same time, the functional form significantly worsen the reliability of most transfers involving WP and ZP regardless of whether assumed to be a policy or study site. However, transfers from WP to MA and MZ are improved. It is important to note that MA and MZ are provinces with the richest theatrical offers among all (see Table A1) and the highest number of frequent theater-goers (as indicated in Table 3). Thus, it is not surprising that controlling for sociodemographic factors positively influences the reliability of transfers from other provinces to these two provinces.

Interestingly, when looking only at transfers of values associated with maintaining the status quo, i.e., no changes in the current theatrical offer, we can observe that TEs are very low (see Figure A2). The range for the unadjusted unit value approach is from 13% to 170%, and for the income-adjusted unit value approach, it is from 5% to 168%. Assuming a TE of 140% as a reliability threshold, transferring value associated with maintaining the status quo with the use of the unit value approach can be considered valid in 29 out of 30 cases, and 28 out of 30 cases after income-adjustment. The functional form again hampers transfers involving WP and ZP but distinguishes the central-east Poland cluster, including LU, LD, MA, and MZ, with TEs within this cluster ranging from 0% to 44%.

6 Conclusions

In this study, we investigate the accuracy of BT, one of the non-market valuation methods, when applied to transferring marginal values for an experienced cultural good, such as theater performances. We examine the BT accuracy based on value estimates derived with the use of identical valuation methods and in the same time period, which helps us provide a pure investigation of the accuracy of transfers over space, excluding confounding factors and other dimensions of possible variation (such as temporal or methodological differences). We argue that assessments such as these provide a novel input to the literature on non-market valuation in cultural economics and can help build a weight of evidence regarding the validity and reliability of transfer methods for cultural goods.

The main findings from our study look promising for the future use of BT for economic valuation of cultural goods. The accuracy of our BTs, as measured by common indicators of validity and reliability, is similar to the accuracy of BTs reported in other areas where the valuation method is broadly used, such as environmental economics (Johnston et al. 2015a).

Based on the validity assessments, 36% of the unadjusted unit value BTs and 32% of the income-adjusted unit value BTs are found invalid, while the percentage for the function BTs is 57%. In turn, in a review of a range of existing BT studies predominantly in environmental and resource economics, Rosenberger (2015) finds that the average rate of invalid BTs is 63%. This shows that for our empirical data on cultural goods and for the better-performing unit value BTs, the validity of value estimates is more often confirmed than rejected, pointing to a better average validity than the one identified by Rosenberger (2015) for environmental goods.

Regarding the reliability, our mean TEs for the unadjusted and income-adjusted unit value transfers are about 180%, while their counterpart for the function transfer is equal to 315%. These magnitudes are larger than the mean TEs observed by in the review of studies by Rosenberger (2015). Specifically, Rosenberger (2015) finds the mean TEs to be 140% for unit value transfers and 65% for function transfers. Although our function transfers do not generate TEs of a comparable mean to those found for environmental goods, our unit value transfers perform relatively comparably in terms of reliability.

The choice of the BT approach may have significant implications for the transfer's validity and reliability. In our study, unit value transfers (including both unadjusted and income-adjusted transfers) perform, on average, better than function transfers in terms of both validity and reliability of the transfers. While this stays in contrast to the popular conclusion from the

environmental economics literature that function transfers perform better (e.g., Johnston and Rosenberger 2010), this is in line with the existing BT accuracy assessments in cultural economics. Most of the BT accuracy studies in culture recommended the use of unit value transfer approaches (Mourato et al. 2014; Fujiwara et al. 2018; Lawton et al. 2018), and our findings bring a new weight of evidence supporting this suggestion.

The accuracy of BT in performing arts may also significantly hinge on the type of performance, for which the BT is conducted. Our results provide some tentative evidence that performance types that are more known in the population can be related to transfers of better validity and reliability. Specifically, based on the empirical data, we argue that, on average, people are more familiar and have a broader experience with watching entertainment and drama performances than children's and experimental performances. We observe the former two to be related to slightly better validity and reliability, however, we note that the evidence is not strong and more research exploring this relationship is needed.

Although the accuracy of BT for a cultural good, such as performing arts, does not fall behind the BT accuracy in other areas, an intriguing question is why the method has not been applied more broadly in culture. We identify three main reasons for this state of (the scarcity of) the BT use. First, there is a relatively small number of primary valuation studies in cultural economics, which could be used as a source of value estimates for BTs (eftec 2005; Tuan et al. 2009; Lawton et al. 2021a). Availability of primary estimates could encourage BTs, as it makes it easier to find study sites matching policy sites in terms of valuation contexts, populations, and other needed dimensions. Second, the literature argues that cultural goods possess unique characteristics and convey a variety of abstract meanings difficult to compare across different locations and populations (Noonan 2003). This makes transferring values of cultural goods a complex task. Third, the existing valuation studies in culture are often site-specific contingent valuation, which hinders transferring values to other contexts.

To make the BT a more accessible valuation method for cultural goods, an obvious response would include that to begin with, more primary research needs to be conducted to provide value estimates for transfers. We argue, however, that it is not only the matter of the number of available primary studies, but also the characteristics of the estimates they deliver. Instead of having more site-specific contingent valuation inquiries, efforts may be directed towards designing valuation studies defining cultural goods in more generic terms. We attempt to undertake this path by using theater performance types in the DCE.

When using DCE-based value estimates for BT, a fundamental question is whether transfers of marginal values are justified. When transferring value estimates, it is critical to consider diminishing marginal utility and variations in relative scarcity across sites. Richardson et al. (2015) argue that in BTs, study-site values cannot be simply multiplied by ratios controlling for differences in the baseline amounts of the evaluated goods between the policy and study sites, but one needs to take into account the entire context of the primary valuation study and ensure a match with the policy site. Because of these considerations, our transfers are based on goods measured relative (i.e., percentage) terms. Johnston and Zawojska (2020) outline a theoretical model, supported by an empirical application, that shows that while transfers of values for goods measured in cardinal units impose a potentially strong assumption of constant marginal utility, transfers of values for goods measured in relative units accommodate diminishing marginal utility. Hence, BTs for goods defined in relative units appear particularly helpful when non-negligible differences exist in baseline conditions between sites.

Overall, our results suggest that using estimates from DCE studies may offer a promising approach for BT applications in culture. Given the complex nature of cultural goods, extrapolating values of particular attributes rather than a single undecomposed value may reduce a risk of BT providing inaccurate estimates for a policy site. We believe that this analysis can help guide future use of the BT method when conducted for experienced cultural goods, such as performing art pieces.

References

Bateman, I. J., Brouwer, R., Ferrini, S., Schaafsma, M., Barton, D. N., Dubgaard, A., Hasler, B., Hime, S., Liekens, I., Navrud, S., De Nocker, L., Ščeponavičiūtė, R., and Semėnienė, D. (2011). Making benefit transfers work: deriving and testing principles for value transfers for similar and dissimilar sites using a case study of the non-market benefits of water quality improvements across Europe. Environmental and Resource Economics, 50(3), 365-387.

Barton, D. N. (2002). The transferability of benefit transfer: contingent valuation of water quality improvements in Costa Rica. Ecological Economics, 42(1-2), 147-164.

Bartosiewicz, S., and Stańczyk, E. (2014). Kontynuacja rozważań na temat wybranych aspektów sytuacji społeczno-gospodarczej ściany wschodniej w porównaniu z resztą Polski w latach 2003–2011. Studia i Prace Wydziału Nauk Ekonomicznych i Zarządzania 36(2), 159-177.

Bishop, R. C., and Boyle, K. J. (2019). Reliability and validity in nonmarket valuation. Environmental and Resource Economics, 72, 559-582.

Boyle, K. J., and Bergstrom, J. C. (1992). Benefit transfer studies: myths, pragmatism, and idealism. Water resources research, 28(3), 657-663.

Brito, P., and Barros, C. (2005). Learning-by-consuming and the dynamics of the demand and prices of cultural goods. Journal of cultural economics, 29, 83-106.

Brouwer, R., and Navrud, S. (2015). The use and development of benefit transfer in Europe. In R. J. Johnston, J. Rolfe, R. S. Rosenberger, and R. Brouwer (Eds.), Benefit Transfer of Environmental and Resource Values: A Guide for Researchers and Practitioners (pp. 71-83). Dordrecht, The Netherlands: Springer.

Brown, J. (2004). Economic Values and Cultural Heritage Conservation: Assessing the Use of Stated Preference Techniques for Measuring Changes in Visitor Welfare. Doctoral dissertation, Imperial College London.

Browning, E., Christie, M., Czajkowski, M., Chalak, A., Drummond, R., Hanley, N., Jones, K. E., Kuyer, J., and Provins, A. (2024). Valuing the economic benefits of species recovery programmes. People and Nature, 6(2), 894-905.

Carson, R. T., and Louviere, J. J. (2011). A common nomenclature for stated preference elicitation approaches. Environmental and resource economics, 49, 539-559.

Castiglione, C., and Infante, D. (2016). Rational addiction and cultural goods: the case of the Italian theatregoer. Journal of Cultural Economics, 40, 163-190.

Czajkowski, M., Bylicki, M., Budziński, W., and Buczyński, M. (2022). Valuing externalities of outdoor advertising in an urban setting–the case of Warsaw. Journal of Urban Economics, 130, 103452.

Czajkowski, M., and Ščasný, M. (2010). Study on benefit transfer in an international setting. How to improve welfare estimates in the case of the countries' income heterogeneity? Ecological Economics, 69(12), 2409-2416.

Desvousges, W. H., Naughton, M. C., and Parsons, G. R. (1992). Benefit transfer: conceptual problems in estimating water quality benefits using existing studies. Water Resources Research, 28(3), 675-683.

eftec (2005) Valuation of the Historic Environment - the scope for using results of valuation studies in the appraisal and assessment of heritage-related projects and programmes. Final Report. Report to English Heritage, the Heritage Lottery Fund, the Department for Culture, Media and Sport and the Department for Transport. https://historicengland.org.uk/images-books/publications/valuation-historic-environment/valuation-historic-environment-final-rep/ (Accessed July 31, 2023).

Freeman, A. M. (1984). On the tactics of benefit estimation under Executive Order. In V. K. Smith (Ed.), Environmental Policy Under Reagan's Executive Order: The Role of Benefit-Cost Analysis (pp. 167-186). Chapel Hill: University of North Carolina Press.

Frey, B. S. (1999). State support and creativity in the arts: Some new considerations. Journal of cultural Economics, 23, 71-85.

Fujiwara, D., Bakhshi, H., Mourato, S., Lawton, R., Hotopp, U., Lagarde, A., and Davies, J. (2018). The Economic Value of Culture: A Benefit Transfer Study. Report to the Department for Digital, Culture, Media and Sport. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/745780/The_Economic_Value_of_Culture_-_A_Benefit_Transfer_Study_-_Final_report.pdf (Accessed October 7, 2022).

Hanley, N., Colombo, S., Tinch, D., Black, A., and Aftab, A. (2006). Estimating the benefits of water quality improvements under the Water Framework Directive: are benefits transferable?. European Review of Agricultural Economics, 33(3), 391-413.

Hanley, N., and Czajkowski, M. (2019). The role of stated preference valuation methods in understanding choices and informing policy. Review of Environmental Economics and Policy, 13(2), 248-266.

Hansen, T. B. (1997). The willingness-to-pay for the Royal Theatre in Copenhagen as a public good. Journal of cultural economics, 21, 1-28.

Johnston, R. J., Boyle, K. J., Adamowicz, W., Bennett, J., Brouwer, R., Cameron, T. A., Hanemann, W. M., Hanley, N., Ryan, M., Scarpa, R., Tourangeau, R., and Vossler, C. A. (2017). Contemporary guidance for stated preference studies. Journal of the Association of Environmental and Resource Economists, 4(2), 319-405.

Johnston, R. J., Boyle, K. J., Loureiro, M. L., Navrud, S. and Rolfe, J. (2021). Guidance to enhance the validity and credibility of environmental benefit transfers. Environmental and Resource Economics, 79(3), 575-624.

Johnston, R. J., Rolfe, J., and Zawojska, E. (2018). Benefit transfer of environmental and resource values: progress, prospects and challenges. International Review of Environmental and Resource Economics, 12(2-3), 177-266.

Johnston, R. J., and Rosenberger, R. S. (2010). Methods, trends and controversies in contemporary benefit transfer. Journal of Economic Surveys, 24(3), 479-510.

Johnston, R. J., Rolfe, J., Rosenberger, R. S., and Brouwer, R. (2015a). Benefit Transfer of Environmental and Resource Values. A Guide for Researchers and Practitioners. Dordrecht, The Netherlands: Springer.

Johnston, R. J., Rolfe, J., Rosenberger, R. S., and Brouwer, R. (2015b). Introduction to benefit transfer methods. In R. J. Johnston, J. Rolfe, R. S. Rosenberger and R. Brouwer (Eds.), Benefit transfer of environmental and resource values: a guide for researchers and practitioners (pp. 19-59). Dordrecht, The Netherlands: Springer.

Kaul, S., Boyle, K. J., Kuminoff, N. V., Parmeter, C. F., and Pope, J. C. (2013). What can we learn from benefit transfer errors? Evidence from 20 years of research on convergent validity. Journal of Environmental Economics and Management, 66(1), 90-104.

Kolasa-Nowak, A. (2011). Eastern Poland and sociological analyses of regional differentiation, Annales. Universitatis Maria Curie-Skłodowska, 36(2), 71-84.

Lawton, R., Fujiwara, D., Mourato, S., Bakhshi, H., Lagarde, A., and Davies, J. (2018). The Economic Value of Culture: A Benefit Transfer Study. Nesta. https://media.nesta.org.uk/documents/Cathedrals_and_Historic_Cities_report_Nesta_and_Simetrica_021018.pdf (Accessed October 7, 2022).

Lawton, R., Fujiwara, D., Arber, M., Maguire, H., Malde, J., O'Donovan, P., Lyons, A., and Atkinson, G. (2021a). DCMS Rapid Evidence Assessment: Culture and Heritage Valuation Studies - Technical Report. Simetrica Jacobs. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/955142/REA_culture_heritage_value_Simetrica.pdf (Accessed July 23, 2023).

Lawton R., Fujiwara, D., Arber, M., Radosevic, D., Lagarde, A., O'Donovan, P., Davies J., Bakhshi H. (2021b). Arts Council England: Regional Galleries and Theatres Benefit Transfer Report. Simetrica Jacobs. https://www.artscouncil.org.uk/sites/default/files/download-file/Arts%20Council%20England%20-%20Regional%20Galleries%20and%20Theatres%20Benefit%20Transfer%20Report.pdf (Accessed June 1, 2023).

Lawton, R., Fujiwara, D., Mourato, S., Bakhshi, H., Lagarde, A., and Davies, J. (2021c). The economic value of heritage in England: A benefit transfer study. City, Culture and Society, 26, 100417.

Lindhjem, H., and Navrud, S. (2015). Reliability of meta-analytic benefit transfers of international value of statistical life estimates: tests and illustrations. In R. J. Johnston, J. Rolfe, R.S. Rosenberger and R. Brouwer (Eds.), Benefit Transfer of Environmental and Resource Values: A Guide for Researchers and Practitioners (pp. 441-464). Dodrecht, The Netherlands: Springer.

Loomis, J. B. (1992). The evolution of a more rigorous approach to benefit transfer: benefit function transfer. Water Resources Research, 28(3), 701-705.

Loomis, J. B. (2015). The use of benefit transfer in the United States. In R. J. Johnston, J. Rolfe, R.S. Rosenberger and R. Brouwer (Eds.), Benefit Transfer of Environmental and Resource Values: A Guide for Researchers and Practitioners (pp. 441-464). Dodrecht, The Netherlands: Springer.

Mahieu, P. A., Andersson, H., Beaumais, O., Crastes dit Sourd, R., Hess, S., and Wolff, F. C. (2017). Stated preferences: a unique database composed of 1657 recent published articles in journals related to agriculture, environment, or health. Review of Agricultural, Food and Environmental Studies, 98, 201-220.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behaviour. In P. Zarembka (Ed.), Frontiers in Econometrics (pp. 105-142). New York: Academic Press.

Mourato, S., Fimereli, E., Contu, D., Gaskell, C., and Boniatti-Pavese, C. (2014). The economic benefits of cultural built heritage interiors conservation from climate change damages in Europe: Final Report. Grantham Research Institute on Climate Change and the Environment. https://www.climateforculture.eu/index.php?inhalt=downloadandfile=pages/user/downloads/project_results/D_06.1_final_publish.pdf (Accessed July 23, 2023).

Navrud, S., and Ready, R. (2007). Review of methods for value transfer. In S. Navrud and R. Ready (Eds.), Environmental value transfer: Issues and methods (pp. 1-10). Dordrecht: Springer Netherlands.

Newbold, S., David Simpson, R., Matthew Massey, D., Heberling, M. T., Wheeler, W., Corona, J., and Hewitt, J. (2018). Benefit transfer challenges: perspectives from US practitioners. Environmental and Resource Economics, 69, 467-481.

Noonan, D. S. (2003). Contingent valuation and cultural resources: A meta-analytic review of the literature. Journal of cultural economics, 27, 159-176.

Provins, A., Pearce, D., Ozdemiroglu, E., Mourato, S., and Morse-Jones, S. (2008). Valuation of the historic environment: The scope for using economic valuation evidence in the appraisal of heritage-related projects. Progress in planning, 69(4), 131-175.

Richardson, L., Loomis, J., Kroeger, T., and Casey, F. (2015). The role of benefit transfer in ecosystem service valuation. Ecological economics, 115, 51-58.

Rolfe, J., Johnston, R. J., Rosenberger, R. S., and Brouwer, R. (2015). Introduction to benefit transfer methods. In R. J. Johnston, J. Rolfe, R. S. Rosenberger and R. Brouwer (Eds.), Benefit Transfer of Environmental and Resource Values: A Guide for Researchers and Practitioners. Dodrecht, The Netherlands: Springer.

Rolfe, J., Bennett, J., and Kerr, G. (2015). Applied benefit transfer: an Australian and New Zealand policy perspective'. In R. J. Johnston, J. Rolfe, R.S. Rosenberger and R. Brouwer (Eds.), Benefit Transfer of Environmental and Resource Values: A Guide for Researchers and Practitioners. Dodrecht, The Netherlands: Springer.

Rosenberger, R. S., and Loomis, J. B. (2017). Benefit transfer. In K. J. Champ, K. Boyle and T. C. Brown (Eds.) A Primer on Nonmarket Valuation (2nd ed., p.431-462). Dordrecht, The Netherlands: Springer.

Rosenberger, R. S., and Stanley, T. D. (2006). Measurement, generalization, and publication: Sources of error in benefit transfers and their management. Ecological economics, 60(2), 372-378.

Scarpa, R., Thiene, M., and Train, K. (2008). Utility in willingness to pay space: a tool to address confounding random scale effects in destination choice to the Alps. American journal of agricultural economics, 90(4), 994-1010.

Seaman, B. A. (2006). Empirical studies of demand for performing arts. In D. Throsby and V. Ginsburgh (Eds.), Handbook of the economics of art and culture (pp. 415–472). Amsterdam: North Holland.

Smolny, K. (2017). The idea of availability as basis of subsidising public theaters in Poland. Finanse, Rynki Finansowe, Ubezpieczenia, 89(1), 389-400.

Statistics Poland (2018). Area and population in the territorial profile in 2018. Warsaw, Poland: Statistics Poland, Standards and Registers Department. https://stat.gov.pl/download/gfx/portalinformacyjny/pl/defaultaktualnosci/5468/7/15/1/powierzchnia_i_ludnosc_w_przekroju_terytorialnym_w_2018.pdf (Accessed July 28, 2023).

Statistics Poland (2019a). Human capital in Poland in the years 2014-2018. Warsaw, Gdańsk, Poland: Statistics Poland. fromhttps://stat.gov.pl/files/gfx/portalinformacyjny/pl/defaultaktualnosci/5501/8/7/1/kapital_ludzki_w_polsce_w_latach_2014-2018.pdf (Accessed July 28, 2023).

Statistics Poland (2019b). Statistical Yearbook of the Regions – Poland 2018. Statistical Publishing Establishment: Warsaw, Poland.

Statistics Poland (2020). Participation in culture 2019. Warsaw, Poland: Statistical Publishing Establishment. https://stat.gov.pl/download/gfx/portalinformacyjny/en/defaultaktualnosci/3310/2/2/1/participation_in_culture_in_2019.pdf (Accessed July 31, 2023).

Statistics Poland (2023). Local Data Bank. https://bdl.stat.gov.pl/bdl/dane (Accessed July 27, 2023).

Termansen, M., McClean, C. J., and Jensen, F. S. (2013). Modelling and mapping spatial heterogeneity in forest recreation services. Ecological economics, 92, 48-57.

Theatre Institute (2019). Theatre in Poland 2019. Documentation of the season 2017/2018. Warsaw, Poland: Theatre Institute.

Throsby, C. D. (1990). Perception of Quality in Demand for the Theatre. Journal of cultural economics, 14, 65-82.

Throsby, C. D., and Withers, G. A. (1979). The economics of the performing arts. New York: St. Martin's Press.

Towse, R. (2010). A textbook of cultural economics. Cambridge: Cambridge University Press.

Train, K. E., and Weeks, M. (2005). Discrete choice models in preference space and willingness-to-pay space. In R. Scarpa and A. Alberini (Eds.), Applications of Simulation Methods in Environmental and Resource Economics (pp. 1–16). Dodrecht, The Netherlands: Springer.

Tuan, T. H., Seenprachawong, U., and Navrud, S. (2009). Comparing cultural heritage values in South East Asia–Possibilities and difficulties in cross-country transfers of economic values. Journal of cultural heritage, 10(1), 9-21.

Wiśniewska, A. (2019). Quality attributes in the non-market stated-preference based valuation of cultural goods. Central European Economic Journal, 6(53), 133-150.

Wiśniewska, A., and Czajkowski, M. (2019). Designing a socially efficient cultural policy: The case of municipal theaters in Warsaw. International Journal of Cultural Policy, 25(4), 445-457.

Zieba, M. (2009). Full-income and price elasticities of demand for German public theatre. Journal of Cultural Economics, 33, 85-108.

Zieba, M. (2011). An analysis of technical efficiency and efficiency factors for Austrian and Swiss non-profit theatres. Swiss Journal of Economics and Statistics, 147(2), 233-274.

Appendix

Table A1. Statistics for the studied provinces for 2018.

	Local public expenditures on theaters per capita per year [PLN]*	**Number of public theater venues****	**Number of performances in public theaters****	**Residents per seat in theaters and music institutions**
LU	10	5	704	1,142
LD	28	9	2,420	492
MA	29	20	4,777	420
MZ	39	27	10,341	269
WP	24	11	2,122	616
ZP	21	6	1,516	382
Averages for the six provinces	25	13	3,647	554
Averages for Poland	27	8.75	2,239	534

Notes: The data comes from Statistics Poland (2023).
* It includes expenditures of all (three levels of) local governments in Poland: municipalities, counties, and provinces. It does not include the central government expenditures spent in a province.
** We acknowledge that due to inconsistencies in reporting statistics for theaters, these numbers are for 121 public theaters and 19 theaters. There are no available statistics for the 121 public theaters only.

	Model for LU	Model for LD	Model for MA	Model for MZ	Model for WP	Model for ZP
Mean effects:						
Status quo	0.33***	0.24***	0.29***	0.19***	0.15	0.12***
	(0.04)	(0.09)	(0.01)	(0.00)	(0.12)	(0.03)
Entertainment	-0.04	0.35***	0.19***	0.09***	0.42*	0.05
	(0.06)	(0.13)	(0.05)	(0.01)	(0.23)	(0.05)
Drama	0.16***	0.17	0.06	0.19***	-0.25	-0.08
	(0.05)	(0.14)	(0.05)	(0.01)	(0.23)	(0.08)
Children's	0.21***	-0.02	-0.37***	0.26***	-0.09	-0.11**
	(0.06)	(0.12)	(0.06)	(0.01)	(0.21)	(0.05)
Experimental	0.12	0.10	0.07	-0.11***	-0.38*	-0.14**
	(0.09)	(0.14)	(0.05)	(0.01)	(0.21)	(0.07)
Cost (100 PLN)	-3.55***	-2.99***	-3.91**	-4.00**	-2.10***	-4.41***
	(1.13)	(0.96)	(1.53)	(1.76)	(0.58)	(1.66)
Standard deviations:						
Status quo	0.81***	0.61***	0.77***	0.56***	0.92***	0.63***
	(0.03)	(0.12)	(0.02)	(0.01)	(0.20)	(0.03)
Entertainment	1.21***	0.33***	0.09***	0.47***	0.97***	0.66***
	(0.06)	(0.14)	(0.03)	(0.01)	(0.23)	(0.04)
Drama	0.77***	0.57***	0.27***	0.50***	0.71***	0.52***
	(0.03)	(0.09)	(0.02)	(0.01)	(0.22)	(0.07)
Children's	0.85***	0.43***	0.65***	0.73***	1.48***	0.29***
	(0.05)	(0.13)	(0.05)	(0.01)	(0.32)	(0.06)
Experimental	1.15***	0.50***	1.16***	0.66***	1.13***	0.25***
	(0.05)	(0.10)	(0.03)	(0.01)	(0.25)	(0.03)
Cost (100 PLN)	3.94***	2.69***	4.97***	3.90***	1.86***	4.53***
	(1.10)	(0.82)	(1.58)	(1.48)	(0.57)	(1.26)
Log-likelihood at convergence	-290.54	-343.86	-458.09	-820.33	-446.15	-185.76
Log-likelihood at constant(s) only	-377.86	-510.93	-615.00	-1084.94	-581.56	-276.74
AIC/n	1.10	0.94	1.02	1.02	1.07	0.96
BIC/n	1.31	1.10	1.16	1.11	1.21	1.21
Number of observations (n)	576	788	952	1664	888	444
Number of respondents	144	197	238	416	222	111

Table A2. Results of the basic MXL models in WTP space (for unit value transfers).

Notes: Standard errors are given in brackets. ***, **, and * represent significance at the 1%, 5%, and 10% level, respectively. The variables on main effects are modeled as random and normally distributed, except for cost, which is assumed to follow a lognormal distribution and the estimates of the underlying, preference-space equivalent, normal distribution are provided. To ensure the model convergence, the cost is divided by 100.

Table A3. Results of the extended MXL models with interactions in WTP space (for function transfers).

	Model for LU	Model for LD	Model for MA	Model for MZ	Model for WP	Model for ZP
Mean effects:						
Status quo	0.32***	0.37***	0.36***	0.28***	-0.06	0.12***
	(0.01)	(0.01)	(0.00)	(0.08)	(0.12)	(0.02)
Entertainment	0.27***	0.51***	0.06***	0.33***	0.09	-0.04
	(0.05)	(0.03)	(0.01)	(0.12)	(0.18)	(0.04)
Drama	0.10***	0.08*	0.53***	0.34***	-0.36*	0.12**
	(0.02)	(0.04)	(0.01)	(0.13)	(0.21)	(0.05)
Children's	0.36***	0.03	-0.47***	0.40***	0.15	-0.22***
	(0.03)	(0.03)	(0.01)	(0.14)	(0.22)	(0.04)
Experimental	0.38***	0.42***	-0.23***	-0.21	-0.59**	-0.29***
	(0.03)	(0.02)	(0.02)	(0.14)	(0.24)	(0.05)
Cost (100 PLN)	-4.04	-4.75**	-4.01	-3.06***	-2.93***	-4.34**
	(2.47)	(2.09)	(3.40)	(0.73)	(0.98)	(1.90)
Standard deviations:						
Status quo	0.86***	0.61***	0.91***	0.61***	0.93***	0.60***
	(0.00)	(0.01)	(0.00)	(0.09)	(0.22)	(0.02)
Entertainment	1.27***	0.24***	0.62***	0.17*	0.87***	0.60***
	(0.04)	(0.02)	(0.00)	(0.10)	(0.22)	(0.03)
Drama	0.81***	0.46***	0.38***	0.29***	0.48***	0.47***
	(0.02)	(0.01)	(0.01)	(0.09)	(0.14)	(0.05)
Children's	0.97***	0.43***	0.69***	0.85***	1.33***	0.18***
	0.02	(0.02)	(0.01)	(0.14)	(0.26)	(0.03)
Experimental	1.17***	0.41***	1.26***	0.62***	1.06***	0.37***
	0.02	(0.01)	(0.01)	(0.14)	(0.23)	(0.05)
Cost (100 PLN)	6.85**	4.23***	7.02**	2.22***	2.07***	4.87***
	(2.97)	(1.42)	(3.25)	(0.60)	(0.61)	(1.43)
Interactions with the variable controlling for living in the province capital:						
Status quo	-0.19***	-0.18***	-0.12***	-0.06	0.32*	-0.24***
	(0.02)	(0.01)	(0.00)	(0.08)	(0.19)	(0.02)
Entertainment	-0.21***	-0.21***	0.22***	-0.28**	-0.63**	-0.31***
	(0.05)	(0.02)	(0.01)	(0.12)	(0.28)	(0.04)
Drama	0.29***	0.21***	-0.27***	-0.13	0.46*	-0.21***
	(0.04)	(0.02)	(0.01)	(0.12)	(0.26)	(0.04)
Children's	-0.24***	0.01	0.03***	-0.07	0.19	-0.13***
	(0.04)	(0.02)	(0.01)	(0.15)	(0.26)	(0.03)
Experimental	-1.26***	-0.74***	-0.18***	0.04	0.87**	-0.49***
	(0.04)	(0.02)	(0.01)	(0.14)	(0.35)	(0.03)
Cost (100 PLN)	-0.61	-0.54	0.30	0.50	-0.45	-0.29
	(1.26)	(1.08)	(1.52)	(0.42)	(0.71)	(1.59)
Interactions with the variable controlling for the university degree:						
Status quo	0.15***	-0.25***	-0.02***	-0.09	-0.22	-0.07***
	(0.02)	(0.02)	(0.00)	(0.08)	(0.19)	(0.02)
Entertainment	-0.12***	-0.76***	0.26***	0.04	1.13***	0.03
	(0.04)	(0.03)	(0.01)	(0.12)	(0.34)	(0.03)
Drama	-0.06	0.16***	-0.63***	-0.04	-0.09	-0.44***
	(0.04)	(0.03)	(0.01)	(0.12)	(0.29)	(0.03)
Children's	-0.11**	-0.20***	0.28***	-0.25*	-0.53*	0.26***
	(0.04)	(0.02)	(0.01)	(0.15)	(0.31)	(0.03)
Experimental	0.39***	-0.02	0.66***	0.21	-0.57	0.25***
	(0.04)	(0.02)	(0.01)	(0.13)	(0.35)	(0.04)

Cost (100 PLN)	0.89 (1.29)	0.43 (1.17)	-2.76 (3.40)	0.42 (0.41)	0.78 (0.67)	-1.35 (1.47)
Log-likelihood at convergence	-261.80	-285.51	-414.70	-710.71	-367.79	-152.64
Log-likelihood at constant(s) only	-351.28	-439.02	-552.73	-955.93	-511.82	-241.55
AIC/n	1.14	0.97	1.07	1.03	1.05	1.00
BIC/n	1.45	1.23	1.28	1.17	1.28	1.40
Number of observations (n)	528	668	852	1460	776	384
Number of respondents	132	167	213	365	194	96

Notes: Standard errors are given in brackets. ***, **, and * represent significance at the 1%, 5%, and 10% level, respectively. The variables on main effects are modeled as random and normally distributed, except for cost, which is assumed to follow a lognormal distribution and the estimates of the underlying, preference-space equivalent, normal distribution are provided. The interactions are modeled as non-random. To ensure the model convergence, the cost is divided by 100. The numbers of respondents used in the models are smaller than the number of respondents per province because of missing data on some of the interaction variables for some respondents.

Figure A1. Box plots for TEs by DCE characteristics.

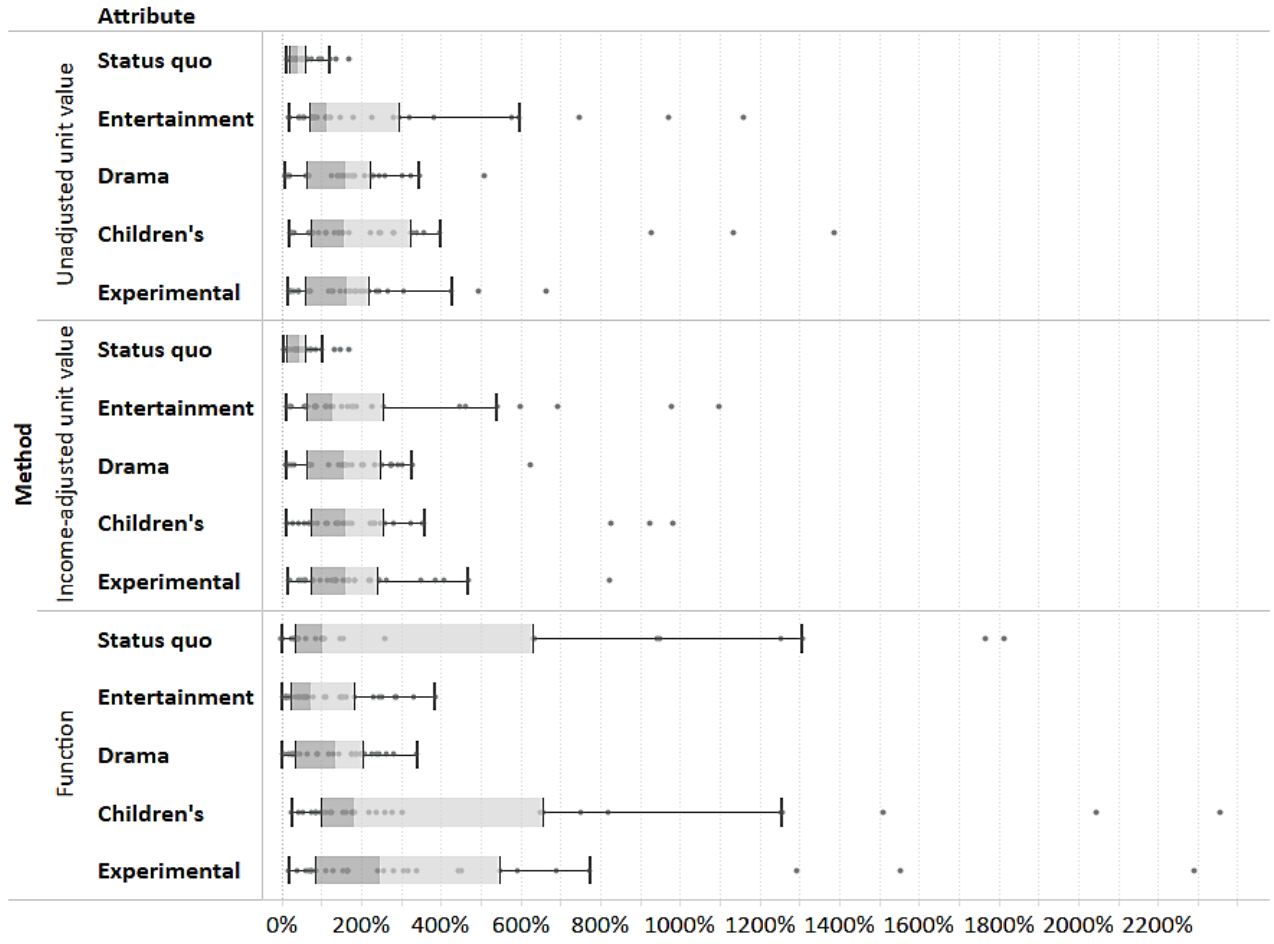


Figure A2. Median TEs for the status quo constant only.

Unadjusted unit value

Policy Site	Study Site LU	LD	MA	MZ	WP	ZP
LU		26	13	43	55	63
LD	35		18	23	39	50
MA	14	15		34	48	58
MZ	74	29	52		21	35
WP	120	64	93	26		18
ZP	170	100	136	55	22	

Income-adjusted unit value

Policy Site	Study Site LU	LD	MA	MZ	WP	ZP
LU		25	36	59	57	63
LD	34		14	46	43	50
MA	56	17		37	33	42
MZ	146	84	58		5	8
WP	134	75	50	5		13
ZP	168	101	72	9	15	

Function

Policy Site	Study Site LU	LD	MA	MZ	WP	ZP
LU		39	0	31	106	102
LD	23		38	2	84	102
MA	1	39		31	106	102
MZ	36	38	44		61	146
WP	1 767	1 255	1 816	1 308		154
ZP	942	636	949	632	257	

University of Warsaw
Faculty of Economic Sciences
44/50 Długa St.
00-241 Warsaw
www.wne.uw.edu.pl
ISSN 2957-0506