Systematic Variation in Waste Site Effects on Residential Property Values: A Meta-Regression Analysis and Benefit Transfer

Marvin Schütt[1] 📓

Abstract

This article presents a meta-analysis based on 727 estimates from 83 hedonic pricing studies to provide new insights on the effects of waste sites on residential property values. Relative to previous meta-analyses on this subject, estimates are corrected for publication bias and the ability of the meta-regression model to produce reliable benefit-transfer estimates is assessed. Proximity to severely contaminated waste sites has a supremely negative impact on residential property values, whereas on average the distance from non-hazardous waste sites has no effect. Correcting for publication bias has a sizeable impact, reducing the average effect size by up to 38%. Benefit-transfer errors based on the meta-regression model are fairly large and, in line with the broader literature, outperform simple value transfer when the underlying data sample is heterogeneous.

Keywords Hedonic price model · Meta-regression analysis · Publication bias · Residential property values · Waste sites

Abbreviations

APE	Average percentage error
BPLM	Breusch and Pagan Lagrangian multiplier
BT	Benefit transfer
FAT	Funnel Asymmetry Test
FEML	Fixed effects multilevel model
MRA	Meta-regression analysis
NPL	National Priority List
OLS	Ordinary least squares
PET	Precision Effect Test
PEESE	Precision Effect Estimate with Standard Error

✉ Marvin Schütt
schuett@economics.uni-kiel.de

[1] Institute for Environmental, Resource and Spatial Economics, Kiel University, Wilhelm-Selig-Platz 1, 24118 Kiel, Germany

PRISMA   Preferred Reporting Items for Systematic Reviews and Meta-Analyses
REML   Random effects multilevel model
WLS   Weighted least squares

1 Introduction

The world's annual generation of waste equalled two billion tonnes in 2016 and is expected to reach 3.4 billion tonnes by 2050 (Kaza et al. 2018). This poses serious threats to the environment in general and may generate externalities for residents living in close proximity to a waste site. These externalities include health risks, offensive smells, noises or unpleasant views (Giusti 2009). However, well-administered waste sites may nullify these adverse effects, indeed they may even be perceived positively due to their employment potential. Accordingly, it is important to understand whether and under what circumstances waste sites significantly affect local residents.

The hedonic pricing framework pioneered by Rosen (1974) is a prominent method used to evaluate the effect of waste sites on local residents.[1] Studies relying on this concept explain the price variation in residential properties as a combination of the value of their respective characteristics, *e.g.,* distance from a waste site. Empirical evidence on this price-distance relationship, however, is ambiguous. The literature displays high variance in its assessment of the magnitude and significance of the effect as well as disagreement on its sign (see *e.g.* Reichert et al. 1992; Du Preez et al. 2016; Poor et al. 2007 and Ready 2010 for landfill effects). Accordingly, it remains unclear whether waste site externalities significantly affect local residents. In response to these open questions, I apply meta-analytic techniques (Stanley and Doucouliagos 2012; Ringquist 2013) to investigate the existence of adverse price effects on proximate residential property values at the aggregate level. The heterogeneity of the empirical results is discussed and explained with reference to the differences in methodological approaches and waste-site characteristics across the literature. Relative to previous meta-analyses on this subject (*e.g.* Braden et al. 2011; Simons and Saginor 2006), estimates are corrected for publication bias and the ability of the meta-regression model to produce reliable benefit-transfer (BT) estimates is assessed. BT can be an especially valuable tool for policymakers in making predictions on effect sizes in areas where time, data, or money constraints do not permit primary studies (Johnston et al. 2015). The present meta-regression analysis (MRA) builds on a meta-sample of 727 observations from 83 studies, going beyond previous MRAs in this area by using 56 studies hitherto unconsidered. The large sample also enables me to add nine moderators to explore new factors explaining heterogeneity. In the framework of cost–benefit analyses, the MRA can support policymakers in making informed decisions on such things as the placement of new waste sites or clean-up activities for hazardous waste sites.

The results confirm that proximity to severely contaminated waste sites has a supremely negative impact on residential property values, whereas on average the distance from non-hazardous waste sites has no effect. Correcting for publication bias has a sizeable impact, reducing the average effect size by up to 38%. The corrected average

---

[1] There are other methods for eliciting this relationship. However, studies that use methods other than hedonic pricing (such as contingent valuation) are not considered here as they rely on different welfare measures, which is likely to lead to problems of incomparability in meta-analyses (Smith and Pattanayak 2002).

effect size translates into a 1.5% to 2.9% property value increase per mile of increased distance from a waste site for a house at the distance of one mile. Together with waste-site and study characteristics, the primary studies' econometric specifications are identified as important dimensions influencing effect size and explaining observed heterogeneity. In particular, the effect size is reduced for cleaned-up waste sites and for residential properties at greater distances from hazardous waste sites, thus reconciling inconsistent previous findings. BT errors based on the meta-regression model are fairly large and, in line with the broader literature, outperform simple value transfer when the underlying data sample is heterogeneous. While the acceptable level of transfer error is context-dependent (Rosenberger 2015; Brander et al. 2006), the predicted levels of transfer error limit practical application accordingly.

The remainder of the paper is organised as follows: The next section presents an overview of results from previous MRAs on this subject, in addition illustrating the requirements for valid and reliable MRA and BT. Section 3 introduces the meta-dataset used in this paper. Section 4 demonstrates the selection of the appropriate model and publication bias control. The results are presented and discussed in Sect. 5. Section 6 concludes.

2 Literature Review

A meta-analysis is commonly described as a statistical analysis of previously reported research findings on a given empirical effect (Stanley and Doucouliagos 2012). In contrast to qualitative reviews, meta-analytic methods can estimate average effect sizes, quantify the extent of variance observed and help explain heterogeneous results (Borenstein et al. 2011). Advances in meta-analytic methodology and the attendant guidelines (Nelson and Kennedy 2009; Stanley et al. 2013; Nelson 2015; Johnston et al. 2018; Stanley and Doucouliagos 2012) have made current meta-analyses more reliable and useful for both academic and practical purposes. More precisely, meta-analyses now typically rely on large meta-sample sizes and substantial sets of moderators to explain heterogeneous results. They control for publication bias and assess the usefulness of meta-regression results for BT applications. Unsurprisingly, meta-analytic tools have enjoyed increasing popularity in economic research over the past decade (Alinaghi and Reed 2018). Within the literature on environmental valuation, recent applications synthesise the empirical evidence pertaining to such things as water-quality improvements (Johnston et al. 2017; Klemick et al. 2018), river restoration (Chen et al. 2019; Brouwer and Sheremet 2017), wetland values (Vedogbeton and Johnston 2020; Chaikumbung et al. 2016) or flood risk (Beltrán et al. 2018).

Within the strand of literature estimating price-distance relationships between waste sites and residential properties, several studies have already analysed and summarised the empirical evidence in a systematic fashion. Table A1 lists the most relevant MRAs, highlighting their findings and main study attributes. The brief discussion in the following section presents the estimated average effect sizes including identified moderators and the type of waste site considered. It also indicates potential limitations. For the MRAs closest to this study (Simons and Saginor 2006; Lipscomb et al. 2013; Braden et al. 2011) I provide a more detailed overview.

2.1 Previous Reviews and Meta-Analyses

The first reviews in this area (Farber 1998; Zeiss 1998; Boyle and Kiel 2001; Jackson 2001; also Brinkley and Leach 2019) were qualitative and aimed to identify moderators explaining the apparent heterogeneity of waste site-related property-price effects (*e.g.,* mean distance from waste site in the primary study, employment opportunities at the waste site, type of waste considered). Later studies confirmed the relevance of some of these research dimensions and added other moderating variables in initial meta-analyses on the topic. They identified the functional form employed, the type of waste site examined and the mean distance from the waste sites as influential research dimensions (Walton et al. 2006; Chèze 2007; Ready 2010). In one of these initial meta-analyses, Ready (2010) reports an average increase in residential property values of 1.3% to 5.9% per mile of increased distance from a landfill depending on the size of the latter. By contrast, Walton et al. (2006) record an average price premium of 6.7% per mile of increased distance for a different set of landfill studies. Chèze (2007) reports an average discount of 3.8% (8.4%) for living one mile closer to non-hazardous (hazardous) landfills or incinerators. Although they contain important initial insights, all of these studies are restricted in value by low sample sizes. Walton et al. (2006) cover 17 estimates from seven studies; Ready (2010) considers 15 estimates from nine landfill studies; Chèze (2007) discusses 12 studies with 45 estimates. This drawback has motivated further research to validate the findings.

One of the largest meta-analyses in this area of research is Simons and Saginor (2006), with 290 observations from 75 articles, 42 of them with 164 observations using the hedonic pricing method. Their meta-dataset comprises studies dealing not only with contaminated waste sites but also with amenities, reliance on surveys, case studies or hedonic regression techniques. They report a 4% mean increase in property value for each mile of increased distance from the respective site (9.5% for studies focusing exclusively on a disamenity). Across specifications, they confirm that the geographic region, the mean distance from the waste site, the type of waste and the announcement of the closure of a site are important moderating dimensions. Although they acknowledge the importance of publication bias, the potential dependence of multiple observations from the same study and differing levels of precision in the estimates, they do not explicitly accommodate these factors in their meta-regressions.

More recently, Lipscomb et al. (2013) have drawn upon 40 studies with 273 observations in their meta-analysis, about 227 of them from 33 articles using the hedonic pricing method (see Table A1 for details). These observations differ not only in the type of disamenity (landfills, hazardous waste sites, power lines, railroad tracks, etc.) or amenity (proximity to water bodies, view, etc.) discussed but also in the methods employed (hedonic regression, travel cost method, contingent valuation, etc.). More than half of the observations are not concerned with price-distance relationships between waste sites and residential properties as the focus here is on an attempt to detect differences in valuation through the choice of elicitation methods. Hence, it deviates from the focus of the present study and will not be discussed any further.[2]

---

[2] There exist two additional studies that are close to this MRA. Both differ in some respects from the focus chosen here, so this review does not consider them in detail. Kiel and Williams (2007) use a probit regression meta-analysis to model the likelihood of being listed as a superfund site and hence do not focus on average effect sizes of superfunds. Saginor et al. (2011) synthesise the empirical evidence of environmental contamination effects on *non*-residential property values.

The study by Braden et al. (2011) is closest in spirit to this meta-analysis. They consider 46 hedonic studies with 129 estimates from various types of waste site, 114 of them from 38 studies of waste-site effects on residential property values. They use weighted least squares (WLS) and ordinary least squares (OLS) techniques with a large set of moderating regressors some of which are new to the relevant literature. They find that a one-mile increase in distance from a terrestrial (aquatic) hazardous waste site leads to an average increase in property values of 3.5% (15.9%). Nuclear and non-hazardous waste sites do not significantly influence property values, with the effects of non-hazardous waste sites being greater in magnitude (3.1% compared to -0.4%). Some of the newly added moderators help to significantly explain the variation in the data. These include control variables indicating the addition of socio-demographic variables in the primary studies' regressions, the use of sales data instead of assessed values and listing on the National Priority List (NPL).[3] To the surprise of the authors, remediating contaminated sites and the mean distance of properties from the waste site do not consistently influence effect sizes for hazardous and non-hazardous waste sites, thus contradicting in the latter case the "fundamental premise of hedonic property valuation of environmental quality" (Braden et al. 2011: 198). One possible explanation for these surprising results may be that they consider observations from discrete and continuous distance specifications simultaneously, also encompassing linear and quadratic distance specifications. These differences in distance definition may lead to incomparability of effect sizes. In addition, the distance-decay effect may only be detectable for *obvious* disamenities like hazardous waste sites. I discuss both aspects in more detail in my remarks on the study selection process in Sect. 3 and in the results section. Despite their methodological improvements over previous meta-analyses, Braden et al. (2011) do not correct for publication bias in their model.[4]

The meta-analyses reviewed provide valuable insights into the likely range of the average effect size and into heterogeneity aspects reflected in *e.g.* the type of waste site considered. However, none of these studies assess their meta-regression results for usage in BT applications or discuss validity and reliability requirements.[5] Similarly, the omission of publication bias controls and the (partly) small sample sizes impose restrictions on their explanatory power. In the present study I emphasise its inherent value for BT, which is also reflected in the study selection criteria. Additionally, I address the potential presence of publication bias in the development of the econometric specification. Further, a large meta-sample in combination with several sets of moderators enable me to assess the robustness of results and conduct subsample analyses. The next section discusses the requirements for consistent and reliable MRA and accurate BT.

---

[3] The most severely contaminated waste sites in the USA are listed on the NPL. These sites have remediation priority (Environmental Protection Agency 2019).

[4] However, they include dummy variables to control for the effect of publication and significance status on the magnitude of the effect size. Whether this serves as a viable approximation is not clear.

[5] There exists, however, some anecdotal evidence. Walton et al. (2006) provide preliminary insights into the potential of BT for landfills but themselves question the validity of this due to the small sample size. Eshet et al. (2007a) assess transferability of results from four individual hedonic estimates of waste-transfer station effects from Eshet et al. (2007b). Again, the small number of estimates restricts external validity. Finally, Braden et al. (2010) use the MRAs from Braden et al. (2011) for BT applications to the Great Lakes Areas of Concern. Due to a different focus in the underlying meta-model, only one variable in the meta-function can be calibrated to distinguish between sites. This limited capacity for differentiating between sites prompts the authors to advise against using their results for out-of-sample predictions.

2.2 Validity and Reliability Requirements

BTs use existing effect-size estimates from one or more previous studies to infer the effect size for a new policy application (Boyle et al. 2013). In principle, BT based on MRA is a form of function transfer, as the meta-equation can be calibrated to fit the new context (Boyle and Wooldridge 2018). Importantly, BT can be a valuable tool in making predictions on effect sizes in areas where time, data or money constraints make primary studies impracticable (Johnston et al. 2015). Consequently, BTs based on MRAs have increasingly been applied in the context of non-market values in recent years, especially for applied cost–benefit analyses (Vedogbeton and Johnston 2020). At the same time, many challenges remain unresolved, casting doubt on the validity and reliability of BT applications under certain conditions (Johnston et al. 2018; Rosenberger 2015; Vedogbeton and Johnston 2020). Arguably, one of the main requirements for a valid combination of studies on both MRA and BT is a "minimal degree of commodity consistency across metadata observations" (Vedogbeton and Johnston 2020: 836). However, the pooling of observations from different studies with different attributes lies at the heart of any MRA. Hence, while commodity consistency is a commonly acknowledged requirement, a too narrow definition of the commodity under consideration can drastically reduce sample size and impose considerable restrictions on statistical analysis (Chaikumbung et al. 2016; Bergstrom and Taylor 2006; Vedogbeton and Johnston 2020). Alongside commodity consistency, welfare-measure consistency and outcome-variable consistency are often called for (Nelson and Kennedy 2009; Vedogbeton and Johnston 2020; Klemick et al. 2018; see also Rosenberger 2015 for a detailed overview).

The trade-off between consistency and sample size is clearly reflected in the meta-analyses reviewed here. One part of the literature (Walton et al. 2006; Chèze 2007; Ready 2010) has synthesised very consistent sets of studies but is limited by small sample sizes restricting a detailed investigation of factors explaining heterogeneity. Other meta-analyses (Simons and Saginor 2006; Braden et al. 2011; Lipscomb et al. 2013) have opted for larger sample sizes and aim to control for greater heterogeneity in their MRA via moderators. However, this happens at the expense of commodity consistency (waste sites combined with power lines or parks, etc.), welfare consistency (hedonic pricing studies combined with studies using contingent valuation or travel cost methods, etc.) or outcome consistency (effects on residential property values combined with effects on non-residential property values) of the pooled observations (see Table A1 for details). In line with the last-named studies in this part of the literature, I argue that *some* diversity in waste sites is needed for insightful MRA (Vedogbeton and Johnston 2020; Nelson 2015). However, as the results of this MRA will also be assessed for BT, I aim for a higher degree of consistency (Bergstrom and Taylor 2006; Rosenberger 2015; Smith and Pattanayak 2002). Hence, outcome validity and welfare validity are consistently upheld, only allowing the inclusion of observations from hedonic pricing studies reporting price-distance relationships for waste sites and residential properties. Similarly, waste sites are the only commodity allowed. The operative definition of waste here is "any substance or object which the holder discards or intends or is required to discard" (European Commission 2008: Article 3). This definition encompasses different types of waste (*e.g.* hazardous or non-hazardous), disposed of at different facilities (*e.g.* landfills or incinerators) and affecting different elements (*e.g.* soil, air or water), thus allowing for a detailed analysis. The study selection criteria introduced in the next section are designed to meet these minimal consistency criteria.

Two observations from the literature motivate the procedure I have chosen to calculate BT errors. First, Boyle and Wooldridge (2018) emphasise that there is no single meta-analytic model necessarily appropriate for both purposes, *i.e.*, explaining heterogeneity and providing low-error BT estimates. Thus, we can hardly expect any *preferred* model with major explanatory power to perform unusually well in terms of transfer error as well (Nelson 2015). However, smaller transfer error-rates may be expectable for subsets of observations that have an even higher degree of commodity consistency (Eshet et al. 2007a), *e.g.*, sharing the severity of pollution. To assess this eventuality, I thus calculate transfer errors based on both the entire meta-dataset and on subsets of more homogeneous studies. This approach also addresses commodity consistency concerns independent of the study-selection criteria (Nelson 2015; Chaikumbung et al. 2016). Second, though it is generally assumed that in terms of transfer error MRA perform better than value transfer (Rosenberger 2015), some studies have shown that this is not necessarily the case (Lindhjem and Navrud 2008; Klemick et al. 2018; Johnston et al. 2018). Intuitively, BT based on MRA is expected to be beneficial in the context of dissimilar sites (Bergstrom and Taylor 2006; Bateman et al. 2011; Johnston et al. 2015), but counterexamples do exist (Rosenberger 2015). Hence, I also calculate simple value transfer errors for both the entire meta-sample and subsets. This approach is conducive to insights in the way transfer errors depend on the degree of commodity consistency associated with the sample and transfer method chosen in this part of the literature.

3 Meta-Dataset

3.1 Selection of Studies

The strategy employed for identifying relevant studies followed the MEAR-Net guidelines[6] for conducting and reporting meta-analyses (Stanley et al. 2013) and involved three steps. First, seven search engines suitable for the complexity of a predefined search query were used to identify initial records.[7] The search query was a combination of synonyms for residential property values, HPM and waste sites (see Table A2 in the supplementary material for a full list). The search resulted in 2,000 initial records. Subsequently, I screened these initial records for eligibility, discarding all spuriously detected studies. Second, for all eligible studies the respective reference lists were checked for additional suitable material. Third, for the resulting record, I searched four previously unused databases (*sciencedirect*, *JSTOR*, *EVRI* and *Google scholar*) manually for studies citing the studies already identified. The second and third steps were repeated until no other relevant studies were found. At the end of this process, 325 studies were included in a preliminary meta-sample (see Figure A1 in the supplementary material for a detailed PRISMA statement). I started the

---

[6] In accordance with the guidelines referred to in Stanley et al. (2013), a second coder separately coded "a substantial proportion" of the final dataset, *i.e.*, 32 studies. Reassuringly, coding ambiguities were attributable to the varying levels of reporting detail in in the primary studies and were reconciled. Coding decisions in the face of ambiguities are commented on in the meta-dataset available in the data repository linked to in the online version of this article. Remaining ambiguity is indicated in the categorisation of moderator variables as 'unclear'.

[7] These were *EBSCO* (including *EconLit*), *ISI Web of Science*, *RePEc*, *opengrey.eu, science.gov, worldwidescience.org* and *wiso-net.de*.

search in April 2018 and finished in December 2018, using the reference management software Citavi (version 5.7) to list the records identified.

For inclusion, the studies had to comply with the following criteria: (1) use of the basic hedonic pricing method, (2) price of residential properties as dependent variable, (3) distance from a waste site as independent variable, and (4) report of all necessary information for standardising the respective regression coefficient and its measure of precision. These restrictions ensure that the studies included measure a *common* effect—fulfilling commodity, welfare and outcome consistency. Despite these study selection criteria, the studies included differ in terms of the waste-related activity they value, *e.g.*, incineration, landfill, smelter. As discussed in the previous section, *some* degree of commodity diversity is inevitable and also necessary for meaningful statistical analysis. However, one might argue that these differences are sufficient to prohibit the combination and ultimately the meta-analysis of these different types of observation.[8] In this study, I prefer to synthesise the studies to explore the effect of heterogeneous types of waste site by means of appropriate control variables. This is in line with (and indeed rather conservative compared to) previous MRAs in this area of research (Simons and Saginor 2006; Lipscomb et al. 2013) but is also in accordance with other MRAs in the field of environmental valuation (Chaikumbung et al. 2016) and with the recommendations of Stanley and Doucouliagos (2012). Further, my hypothesis is that it is not the activity conducted at the waste site that determines the (assumed) disamenity effect. Instead, I conjecture that the potential undesirability of waste sites stems from the type of resultant pollution. More precisely, I consider the hazardousness of the waste site and the element affected (soil, water, air) to be a set of attributes causing the effect of waste sites on residential property values to differ. Importantly, these factors cannot be ascribed to one particular type of waste site alone, *i.e.*, hazardous waste can be both landfilled or incinerated, thus affecting soil or air (see, *e.g.*, Affuso et al. 2010; Zegarac and Muir 1998). Ultimately, this approach conserves the sample size and enables me to examine a wide spectrum of waste-site effects. However, as already indicated, subsample analyses are provided to assess the sensitivity of this decision. More formally, the study selection criteria ensure that studies included in the meta-sample report results from a variant of the following stylized hedonic pricing specification:

$$P = \alpha_0 + \beta_1 \text{DIST} + \sum_{n=2}^{N} \gamma_n X_n + u \tag{1}$$

with P being the residential property value, *DIST* the distance from a waste site, X_n a set of control variables (with β_1 being the estimated coefficient of interest) and u a common error term. Consequently, several sets of studies included in the preliminary dataset of 325 studies had to be discarded. Table A3 lists these studies with the reasons for exclusion.

Two sets of excluded studies are discussed in more detail to emphasise the importance of comparable estimates in the meta-sample. First, studies that use quadratic distance specifications were eliminated because of the absence of information on the estimates' precision. Given the stylised hedonic pricing specification,

$$P = \alpha_0 + \beta_1 \text{DIST} + \beta_2 \text{DIST}^2 + \sum_{n=3}^{N} \gamma_n X_n + u \tag{2}$$

---

[8] I would like to thank one anonymous reviewer for pointing to this important aspect.

with all variables defined as above, the marginal effect of distance is given by.

$$\frac{\partial P}{\partial \text{DIST}} = \beta_1 + 2 * \beta_2 * \text{DIST} \tag{3}$$

While it is frequently possible to calculate the marginal effect from the information given in the studies, the information required to calculate its standard error is usually not provided (in particular, the covariance between the linear and quadratic term is never reported). Gunby et al. (2017) argue convincingly that it is incorrect to include either of the coefficients individually because they represent incomplete information about the marginal effect of distance on the price of properties. Thus, nine studies were discarded. Second, I have excluded 40 studies using discrete distance specifications, *i.e.*, defining the location of a house to be inside or outside a certain radius. Although this is the largest set of excluded studies, omission is unavoidable. First, the respective radius is defined differently across studies, so pooling studies with discrete distance definitions for meta-analytic purposes is only possible if the different measurement units can be matched. This would entail assumptions on the distribution of houses around the respective waste site and on the nature of the distance-decay effect, *e.g.*, assuming that the effect vanishes linearly over distance and that houses are evenly distributed in concentric circles around the waste site (Debrezion et al. 2007). These assumptions would necessarily introduce measurement error unknown both in extent and in nature.[9] Second, discrete distance definitions cannot be matched with continuous measures (Ready 2010) as a continuous effect size like a percentage increase in property prices per mile cannot be aligned with a dummy variable indicating, say, the value reduction of a house situated within a certain radius around a waste site compared to a house outside this radius. Hence, in contrast to Braden et al. (2011), I consider studies with discrete distance specifications to lead to mutually incomparable estimates and exclude them. As a result, the final dataset consists of 83 studies with 727 observations covering 13 countries and spanning approximately 40 years.[10] For a full list of the studies included, see Table A4 in the supplementary material. To facilitate the overview, Table A5 summarises the basic characteristics of the studies included.

The final dataset consists of studies with various model specifications. Accordingly, the standardisation of the respective regression coefficients to a common metric is required to reconcile different distance variable definitions across studies (*e.g.,* feet or kilometres), the functional form used (*e.g.*, linear or logarithmic) and the estimation strategy (*e.g.,* OLS or spatial autocorrelation) (Nelson and Kennedy 2009). In this meta-study, the common effect size selected is the distance elasticity of residential property prices, or more formally:

$$\text{Elasticity} = \frac{\partial \log P}{\partial \log DIST} = \frac{\partial P}{\partial DIST} \times \frac{DIST}{P} \tag{4}$$

---

[9] One reviewer recommended running a separate regression for this set of excluded studies. I agree that it would help strengthen my argument of incomparability and for exclusion. However, in the light of the necessary assumptions needed to convert the respective study estimates into a common effect size, I prefer not to combine these types of study in a separate MRA.

[10] Three observations generally met the inclusion criteria but were still discarded because the reported effect sizes far exceeded reasonable levels, which hints at typing errors or the like. These were two observations from Bilbao-Terol (2009) and the only observation from Li et al. (2015). They reported price-distance elasticities of 92 and 81 (Bilbao-Terol 2009) and 521 (Li et al. 2015). Leaving these observations aside reduces the simple mean of the effect size from 0.991 to 0.042. Attempts to get an explanation for these seemingly bizarre results were unsuccessful.

If original estimates differ from this common effect size, they are consistently converted, see Table A6 in the supplementary material for a detailed description. Elasticities resulting from such conversion are evaluated at the mean of the respective distance and price variables, if applicable. Similarly, the standardisation process for the measures of statistical precision in the original estimates (standard error, t-value or p-value) is summarised in Figure A2 in the supplementary material. The effect size can be interpreted as the percentage change in the price of a residential property in response to a one-percent increase in its distance from a waste site and serves as the dependent variable in this meta-analysis. Given the results from the previous literature, the elasticity is expected to be positive, viz. greater distance from a waste site is expected to be beneficial for residential property values.

3.2 Selection of Moderators

The moderators, their respective definitions and descriptive statistics are summarised in Table 1. The selection of moderators is undertaken on the basis of previous findings as discussed in the literature review.[11] The asterisks (*) indicate previously unconsidered moderators reflecting additional methodological particularities in the primary hedonic pricing studies.

In general, the moderators can be grouped into three categories reflecting site characteristics, data characteristics and researcher decisions on methodology. Considering site characteristics serves to distinguish the effects of different types of waste site, *i.e.*, the severity of contamination, the element affected,[12] the activity status of the site or the continent on which the site is located. Similarly, the number of proximate waste sites and the clean-up stage of hazardous waste sites may influence the estimated effect size. Hazardous waste sites, for example, are expected to have more adverse effects on property values than non-hazardous sites due to their greater expected impact on people's health in their vicinity. Additionally, multiple waste sites are presumably related to greater effect sizes than the single-site case due to the greater likelihood of noise or offensive smells.

Data characteristics reflect some particularities of the respective sample of residential property values, such as mean distance from the waste site, sample size and whether the properties in question were sold rather than assessed. Only a subset of all observations reports the mean distance of houses from a waste site in the respective sample. Two newly introduced moderators serve as potential alternatives. First, a dichotomising moderator indicates studies with reported sample mean distances greater than the mean distance in this meta-sample (4.29 miles). Second, a dummy variable signals whether the definition of the distance variable is in miles or kilometres as opposed to, say, feet or metres. I presume that studies defining their distance variable in miles or kilometres will report smaller estimates than studies with, say, feet or metres as the chosen metric. Full information on both alternatives is available. All else being equal, greater mean distances can be expected to lead to smaller effect sizes. The choice of these dummy moderators is designed to shed

---

[11] In some cases I adopt the naming conventions used by Braden et al. (2011) as I consider them informative and intuitive. I hope that this enhances the comparability of results.

[12] I collected the information for the affected element from the primary studies and the US EPA, where applicable (see https://cumulis.epa.gov/supercpad/cursites/srchsites.cfm). If a site affected more than one element and the main affected element could not be inferred from these sources, the coding was based on the judgements of two separate coders. In ambiguous cases the element was coded as '*unclear*'.

Table 1 Summary statistics of moderators

Moderator	Definition	Mean	SD	Number of observations [studies]
SE*	Standard error of effect size	0.07	0.12	727 [83]
Published	=1 if study is published in a journal, 0 else	0.78	0.41	727 [83]
Year_publish	Year the study is published	2006	7.74	727 [83]
Site characteristics				
Site_m	=1 if multiple sites were present and considered, 0 else	0.39	0.49	727 [83]
HDI*°	HDI for the country of the waste site in the year of the sampled data at subnational level if available, on national level else	0.87	0.04	369 [70]
GDP*°	GDP p.c. in 2010 USD for the country of the waste site in the year of the sampled data	38,308	10,022	469 [81]
NPL				
Not on NPL	=1 if waste site is not on the NPL, 0 else	0.56	0.50	404 [45]
On NPL	=1 if waste site is on the NPL, 0 else	0.25	0.44	184 [22]
Not USA	=1 if waste site is not in the USA and hence not eligible for the NPL, 0 else	0.17	0.37	121 [26]
Unclear	=1 if status on the NPL is unclear, 0 else	0.02	0.16	18 [4]
Activity status				
Inactive	=1 if waste site is inactive, 0 else	0.68	0.47	494 [49]
Active	=1 if waste site is active, 0 else	0.24	0.43	175 [45]
Unclear	=1 if activity status is unclear, 0 else	0.08	0.27	58 [7]
Employment opportunities				
No employment opportunities	=1 if the waste does not offer employment, 0 else	0.31	0.46	226 [44]
Employment opportunities	=1 if the waste offers employment, 0 else	0.25	0.43	180 [41]
Unclear	=1 if employment at the waste site is unclear, 0 else	0.44	0.50	321 [7]
Type of waste site				
Hazardous	=1 if waste is hazardous, 0 else	0.79	0.41	571 [54]
Non-hazardous	=1 if waste is non-hazardous, 0 else	0.17	0.38	125 [30]
Nuclear	=1 if waste is nuclear, 0 else	0.04	0.20	31 [6]

Table 1 (continued)

Moderator	Definition	Mean	SD	Number of observations [studies]
Waste-affected element				
Soil	=1 if the soil is the waste-affected element, 0 else	0.58	0.49	421 [40]
Air	=1 if the air is the waste-affected element, 0 else	0.13	0.34	98 [22]
Water	=1 if a water body is the waste-affected element, 0 else	0.11	0.31	79 [11]
Unclear	=1 if the waste-affected element is unclear, 0 else	0.18	0.38	129 [22]
*Continent**				
North America	=1 if waste site is in North America, 0 else	0.85	0.35	620 [61]
Asia	=1 if waste site is in Asia, 0 else	0.08	0.27	59 [9]
Europe	=1 if waste site is in Europe, 0 else	0.06	0.23	41 [10]
Other	=1 if waste site is on another continent (Africa, Australia), 0 else	0.01	0.10	7 [3]
Clean-up stage				
Not recognised	=1 if waste site is not yet recognised as polluted, 0 else	0.11	0.31	78 [13]
Recognised, no clean-up plan exists	=1 if waste site is recognised as polluted but no remediation plan exists, 0 else	0.37	0.48	267 [31]
Clean-up begun, but not finished	=1 if remediation plan for polluted waste site exists, but waste site is not remediated yet, 0 else	0.16	0.37	115 [20]
Cleaned-up	=1 if polluted waste site is remediated, 0 else	0.06	0.23	42 [11]
Unclear	=1 if clean-up stage is nuclear, 0 else	0.14	0.34	100 [21]
Data characteristics				
Dist_greater_mean*	=1 if mean distance from the waste site is greater than the mean distance in the sample (4.29 miles), 0 else	0.40	0.49	727 [83]
Dist_mean°	Mean distance from the waste site in the sample (in miles)	4.29	7.56	548 [82]
Miles_km*	=1 if the unit of distance is in miles or kilometres, 0 else	0.68	0.47	727 [83]
Sample	Sample size of the study (in 1000 s)	6.02	19.95	727 [83]
Sale_ind	=1 if price is sales price on individual house level, 0 else	0.84	0.37	727 [83]
Methodology: Econometric specification				
Num_sig_var*°	Number of significant variables used ($p < 0.05$)	15.10	9.44	396 [76]

Table 1 (continued)

Moderator	Definition	Mean	SD	Number of observations [studies]
Num_expl	Number of explanatory variables used	21.76	14.19	727 [83]
Oth_disamen	=1 if other disamenities are present and considered, 0 else	0.29	0.45	727 [83]
Oth_amen	=1 if other amenities are present and considered, 0 else	0.29	0.45	727 [83]
Access	=1 if accessibility options (highway, distance to central-business-district) are present and considered	0.46	0.50	727 [83]
Industry	=1 if industrial activity is present and considered, 0 else	0.12	0.33	727 [83]
Demoecon	=1 if socio-economic control variables are included, 0 else	0.75	0.43	727 [83]
Time_control*	=1 if a time control is included (e.g., dummy variable, inflation adjustment), 0 else	0.78	0.41	727 [83]
Direction*	=1 if a direction-related variable is controlled for (prevailing wind / compass direction), 0 else	0.07	0.25	727 [83]
Interaction*	=1 if any interaction terms with the distance variable exist, 0 else	0.10	0.30	727 [83]
Methodology: Estimation strategy				
Log_log	=1 if regression is of double-logarithmic type, 0 else	0.63	0.48	727 [83]
OLS_spatial	=1 if OLS or a spatial model specification is used, 0 else	0.97	0.17	727 [83]

The asterisk (*) denotes moderators that have not been used in this context before. The circle (°) indicates moderators with missing information inducing a reduced number of observations. The categorical variables have mutually exclusive levels, *e.g.*, 'not on NPL' signifies that the NPL status is clear and that the waste site is located in the USA. The number of studies within a category of moderators sometimes exceeds the number of studies in this meta-study. This reflects the strategy of some studies analysing a waste site *e.g.* at different points in time with varying clean-up stages

more light on the surprising absence of a distance-decay effect in the findings by Braden et al. (2011), as discussed in Sect. 2.

Unique methodological approaches in estimation strategy or econometric specification are also suspected of systematically influencing results. Although there is no fixed set of prescribed variables for inclusion in a hedonic regression, estimates from regressions with a very small set of control variables are likely to suffer from misspecification bias. Consequently, the reported effect size may also be biased (Wooldridge 2010; Phaneuf and Requate 2017). As a response to this hazard, I include moderators that control for the number and type of explanatory variables. In the same vein, uncommon functional forms or model specifications such as Box-Cox transformations or inverse distance specifications may influence the effect size. Two control variables address these eventualities.

Finally, three moderators are included that do not belong to the categories referred to earlier. First, a dummy variable for the publication year controls for time-trend effects. Two additional moderators aim to address publication bias, potentially influencing the effect sizes assembled. Publication bias occurs when the selection of results by a researcher or the selection of studies by a journal are dictated by statistical significance or theoretical expectations (Stanley and Doucouliagos 2012). Accordingly, if publication bias is present, there will be larger and more significant findings in the accessible literature that do not reflect the true population parameter (Card and Little 2016). Hence, there is good reason to believe that the standard error of the effect size is positively correlated with the effect size, which leads to its inclusion as a moderating variable. Additionally, the peer-review process itself may introduce changes in the set and composition of reported findings and this may also affect the effect size. Hence, a dummy moderator indicates estimates from studies published in peer-reviewed journals to control for this possibility. In Sect. 4 the potential presence of publication bias is reflected in the development of the econometric specification.

3.3 Summary Statistics

The distribution of the effect size and its dependence on the conditions prevailing in the respective studies are of primary interest in this study. Figure A3 and Table 2 illustrate these two aspects. In Figure A3 the effect size is depicted in the form of a frequency distribution. With 194 of 727 estimates being negative, there is a tendency towards positive elasticity values in the meta-sample. However, approximately 70% of the estimates are between -0.1 and 0.1. This clearly indicates that the majority of estimated elasticities are clustered around zero. With only four observations greater than 1 in absolute terms, the price-distance relationship under investigation can be summarised as inelastic in almost all cases. In addition to the overall distribution of the effect size shown in Figure A3, a more nuanced picture may provide an initial impression of the heterogeneity observed. For this purpose, the summary statistics of the effect size are illustrated in Table 2, where the unweighted mean, fifth and 95th percentiles are displayed for the whole sample and several sets of sub-samples defined by selected site characteristics.

Table 2 indicates that the mean effect size is positive for the whole sample and in most of the subsamples. The magnitude of the effect lies in the range of the results of the MRAs discussed earlier. Bearing in mind the definition of the effect size as an elasticity, the unweighted mean effect size for a house one mile away from a waste site is a 4.2% increase in property value per mile increase in distance. Turning to the percentiles reveals major disparities in the observations. Observations at the fifth percentile are generally negative, whereas observations at the 95th percentile have consistently positive effect sizes.

Moreover, the effect size from observations at the 95th percentile is approximately four to ten times larger than the mean, depending on the subsample considered. Comparing the effect size by study characteristics provides additional preliminary insights. Published studies apparently show a higher mean effect size than their unpublished counterparts. Cleaning up a contaminated site would seem to be beneficial for residential property values.

The purpose of inspecting summary statistics is to explore the data and to identify tendencies rather than to draw inferences. As Stanley and Doucouliagos (2012) emphasise, simple average effect sizes (weighted or unweighted) are distorted in the presence of publication bias because in that case the meta-sample would not be drawn at random from the underlying population. Additionally, using simple averages implicitly assumes that all observations are treated equally and ignores the potential interdependence of multiple observations per study, heterogeneity across studies and differences in statistical precision. These potential limitations motivate the choice of meta-analytic model(s) in the following section.

4 Methodology

The choice of the appropriate meta-analytic model is a point of ongoing discussion in the literature (Nelson and Kennedy 2009; Stanley and Doucouliagos 2012; Ringquist 2013). The core of the debate revolves around the best identification of, and correction for, publication bias and the justification for either *random-* or *fixed-effects* models (Stanley and Doucouliagos 2017; Alinaghi and Reed 2018). Here, I draw upon the variety of meta-analytic models and consider them an opportunity for ample robustness checks on the results. Accordingly, I begin with a brief overview of meta-model candidates discussed in the literature, and continue with an assessment of publication bias in the meta-dataset. The overview concentrates on the controversies related to *fixed-* and *random-effects* models, interdependence of observations and heteroscedastic error terms. It follows the decision pathways presented in more detail by Feld and Heckemeyer (2011) and by Stanley and Doucouliagos (2012).

4.1 Choice of the Meta-Analytic Model

Multivariate meta-analytic models including moderating variables have become a standard framework to help explain the very likely presence of heterogeneity in effect sizes in applied economic research (Stanley and Doucouliagos 2012; Ringquist 2013).[13] Accordingly, I adopt a general multivariate model framework as a starting point, *i.e.,*

$$\text{Elasticity}_i = \alpha_0 + \sum_{k=1}^{K} \alpha_k C_{k,i} + \varepsilon_i, \quad i = 1, 2, ..., M \tag{5}$$

---

[13] Simple meta-analytic models omitting heterogeneity-explaining moderators have been shown to be biased in the presence of publication selection (Stanley and Doucouliagos 2012). Later, in Sect. 4.2., a control for publication selection (FAT-PET-PEESE) is introduced, proceeding from this naïve framework to calculate corrected mean effect sizes (see Borenstein et al. 2011 and Ringquist 2013 for in-depth discussions of meta-analytic models).

Table 2 Distribution of effect size by selected study characteristics

Study Characteristics	Number of observations [studies]	Mean	P5	P95
Unweighted mean	727 [83]	0.042	−0.183	0.327
Type of Waste Site				
Hazardous	571 [56]	0.042	−0.177	0.327
Non-hazardous	125 [28]	0.038	−0.135	0.206
Nuclear	31 [6]	0.055	−0.343	0.404
Waste-affected element				
Soil	421 [40]	0.036	−0.295	0.387
Air	98 [22]	0.079	−0.004	0.269
Water	79 [11]	0.048	−0.136	0.318
Unclear	129 [22]	0.031	−0.011	0.108
Continent				
North America	620 [61]	0.039	−0.186	0.351
Asia	59 [9]	0.046	−0.295	0.244
Europe	41 [10]	0.075	−0.082	0.265
Other	7 [3]	0.071	0.030	0.084
Clean-up Stage				
Not recognised	78 [12]	0.072	−0.140	0.429
Recognised, no clean-up plan exists	267 [31]	0.060	−0.203	0.404
Clean-up begun, but not finished	115 [20]	0.024	−0.267	0.236
Cleaned-up	42 [11]	−0.044	−0.387	0.115
Unclear	100 [21]	0.033	−0.018	0.132
Publication Status				
Published	567 [69]	0.049	−0.190	0.366
Not published	160 [14]	0.016	−0.138	0.157

The number of studies within a category of moderators sometimes exceeds the number of studies in this meta-study. This reflects the strategy of some studies analysing a waste site *e.g.* at different points in time with varying clean-up stages

with *elasticity* being the standardised effect size, $C_{k,i}$ representing the k th study characteristic attributed to estimate i and ε_i a random error term with $\varepsilon_i \sim N(0, \sigma_i^2)$. Here, α_0 is an estimate of the genuine mean effect size conditional on the set of controls $C_{k,i}$, *i.e.,* an estimate of the magnitude and significance of the price effect of waste-site proximity on residential properties. This type of model is commonly referred to as the *fixed effects* model (Ringquist 2013).[14] It crucially assumes that any deviation from the mean that is not explained by the moderator variables is entirely random due to sampling error (Feld and Heckemeyer 2011). By contrast, the *random* (also known as *mixed*) *effects* model introduces a second error term allowing for unobserved heterogeneity across observations, *i.e.,*

---

[14] Note that the terminology of *fixed-* and *random-effects* is different from their usage in panel models. The explanations in this section are intended to draw clear distinctions despite this ambiguity in wording.

$$\text{Elasticity}_i = \alpha_0 + \sum_{k=1}^{K} \alpha_k C_{k,i} + \upsilon_i + \varepsilon_i, \tag{6}$$

with $\upsilon_i \sim iid\left(0, \tau^2\right)$ depicting unobserved heterogeneity. A standard test for assessing the presence of unobserved heterogeneity is Cochrane's Q-test (Borenstein et al. 2011). For these multivariate models, the null hypothesis of the Q-test assumes that all heterogeneity is explained by the moderating variables. If this can be rejected, the *random effects* model is generally favoured over its *fixed effects* counterpart (Feld and Heckemeyer 2011).

With many studies reporting multiple estimates, there is potential dependence among estimates from the same study through the study design, shared methodology or sample reuse (Stanley and Doucouliagos 2012; Penn and Hu 2019). One way to account for non-independent observations is to use panel-econometric techniques. In this type of model, a second study layer explicitly reflects the nested structure of estimates. Accordingly, the *multilevel* or *hierarchical* model is given by

$$\text{Elasticity}_{ij} = \alpha_0 + \sum_{k=1}^{K} \alpha_k C_{k,ij} + \lambda_j + \varepsilon_{ij}, \; j = 1, 2, ..., S, \tag{7}$$

with j indexing the study level. The Breusch and Pagan Lagrangian multiplier (BPLM) test helps in deciding whether a panel-type model is appropriate. If the null hypothesis of no study-level effect is rejected, there are again two modelling options. First, the study-level effect λ can be estimated as an unobserved study-level error term resulting in a *random effects* multilevel model (REML), or else it can be modelled explicitly by replacing λ with study dummies, which is known as a *fixed effects* multilevel model (FEML) (Stanley and Doucouliagos 2012). The REML critically assumes that the unobserved study effect is uncorrelated with all regressors. If there is reason to suspect correlation, the FEML is the appropriate choice. A robust Hausman test serves as decision rule (Feld and Heckemeyer 2011). Alternatively, if the BPLM test does not support a panel-type model, clustered standard errors can be calculated to correct correlated error terms at the study level. This is especially apposite if the number of clusters is high (Nelson 2015). A related approach assigns equal weights per study or equal weights per sample to avoid undue dominance of studies with many estimates over studies reporting only one (Penn and Hu 2019).[15]

Regardless of choice, any model should be estimated with WLS rather than OLS (Feld and Heckemeyer 2011). The meta-dataset includes studies with widely dispersed estimates and corresponding variances that induce heteroscedasticity in the error term(s). Hence, though estimating Eqs. (5), (6) or (7) by OLS would produce unbiased results, the estimates would be inefficient. Relying on WLS ensures efficient estimates of the coefficients (Wooldridge 2010). The employed analytic weights are the reciprocal error-term variances, which vary in accordance with the type of model chosen. In the case of the *fixed effects* model, the variance σ_i^2 is given or transformed information on precision in the regressions of the original studies (see Sect. 3.1). Hence, heteroscedasticity is easily accommodated by

---

[15] Another solution proposed is to evade the problem by using only one observation per primary study such as the 'preferred' estimate or 'best evidence' per study (Nelson 2004). However, this inherent selection has been criticised as arbitrary, as greatly diminishing sample size, discarding valuable information and conserving publication bias (Nelson 2015; Stanley and Doucouliagos 2012). Alternatively, the mean effect size per study could be used (Stanley 2001). While this also greatly reduces sample size and discards valuable information, the resulting average may still better reflect the 'true' effect size.

using analytic weights $w_i = \frac{1}{\sigma_i^2} = \frac{1}{SE_i^2}$, with *SE* being the standard error of each respective estimate.[16] In the case of *random effects* models, the weight changes to $w_i = \frac{1}{\sigma_i^2 + \tau^2} = \frac{1}{SE_i^2 + \tau^2}$ to incorporate the additional variance introduced by unobserved heterogeneity. In contrast to σ_i^2, the additional element of variance τ^2 is not known to the meta-analyst a priori and must be estimated in a first step.[17] In both cases, the weights chosen reflect the precision of the respective estimates and thus give greater weight to more precise estimates.[18] In contrast to the *fixed effects* weight, however, the *random effects* weight is typically more evenly distributed due to the added constant between-study variance.[19]

This setup identifies four classes of models, all of them estimated by WLS. Considering the wide range of notations and terminologies in the literature, I hope to clearly distinguish them by calling the *fixed effects* model in Eq. (5) WLS-FE and its *random effects* counterpart in Eq. (6) WLS-RE (following Alinaghi and Reed 2018). As introduced above, their respective panel-type counterparts originating from Eq. (7) are referred to as FEML and REML to underscore their multilevel nature (following Stanley and Doucouligaos 2012). This structured decision process is the framework used for selecting the most appropriate meta-analytic model. Choice is determined entirely by the meta-dataset at hand. Starting from the general multivariate model, Cochrane's Q-test provides the criterion for choosing either a WLS-FE or WLS-RE model. In the latter case, study-level effects reflecting the non-independence of multiple estimates from the same study can explain unexplained heterogeneity. A BPLM test helps investigating the existence of such study-level effects, in which case a multilevel model is the appropriate choice. Finally, the robust Hausman test indicates the appropriateness of either the REML or FEML model.

4.2 Publication Bias

A visual method commonly used to detect publication bias is the examination of a funnel plot. In a funnel plot, the effect sizes are plotted against their respective standard error. In an ideal setting without publication bias, the distribution of effect sizes from studies with large samples would cluster around the top of the plot (where precision is high), with

---

[16] Stata has several commands that estimate weighted regressions, with the same resulting regression coefficients and corresponding standard errors in all cases. With no statistical guidance for preferring one command to the other (Ringquist 2013), I here use the simple and widely known *regress* command with manually weighted variables. This can be thought of as OLS applied to weighted observations.

[17] This is achieved by using the *metareg* package in Stata, which, for example, relies on residual maximum likelihood estimation.

[18] Alternatively, heteroscedasticity-robust standard errors may be calculated. However, as weighting the estimates also serves to give greater importance to more precise estimates, I prefer using WLS to OLS with heteroscedasticity-robust standard errors.

[19] There are major differences in the magnitudes of the effect sizes and their respective standard errors that eventually serve as weights. In the *fixed effects* framework, this means that some observations receive a lot of weight while others have barely any influence. While in principle this reflects the desired weighting mechanism, here it would lead to some observations dominating others. This could be handled by following Feld and Heckemeyer (2011) and excluding observations with extremely high weights. Here, however, there is no clear cut-off point, as there are many observations with both large and small weights simultaneously. Thus, the precision weight is logarithmised, or $\log(precision) = \log(\frac{1}{SE_i})$, as this transformation resulted in an approximately normal distribution of the precision variable ($p < 0.001$). The subsequent regression results are close to the *random effects* framework, so that I consider the transformation a viable alternative in this case.

estimates from studies with smaller samples (and lower precision, *i.e.,* higher standard errors) spreading down into the bottom area, thus creating an inverted funnel shape (Borenstein et al. 2011). This would reflect the random deviation from the genuine mean effect due to sampling error. If the funnel plot is asymmetric, this may hint at publication selection biasing the results (Ringquist 2013).

Figure A4 shows two funnel plots. The analysis of alternative funnel plots indicates the sensitivity of the visual impression. On the left-hand side, each point represents a single observation from the meta-sample. The points on the right-hand side are study means. Furthermore, the type of waste site examined in each study serves as a label for the points depicted: hazardous, non-hazardous and nuclear-waste sites. I consider waste-site category labels to reflect the possibility that publication selection only occurs for some types of waste site. The standard funnel plot on the left is a rather symmetric, homogenous plot with many precise estimates clustered around the top. Despite the symmetric impression, there is considerable dispersal of estimates at the top, suggesting that the monetary effect of waste sites on residential property values may be moderated by some study characteristics, such as the type of waste site. The funnel plot of study means on the right-hand side does not confirm the impression of a symmetric graph. Clearly, several estimates lie right to the centre, forming an asymmetric plot. Note, however, that taking study means narrows down the scales of the axes so that the two funnel plots cannot be compared directly. Turning to the waste-site labels in the funnel plot on the left-hand side, it becomes evident that studies examining hazardous waste sites report the most widely spread results. However, there is no clear visual evidence of differences in plot symmetry by waste-site labels. The funnel plot on the right-hand side generally supports this impression, with studies examining non-hazardous waste sites also reporting dispersed findings. The funnel plots taken as a whole suggest publication bias in the sample and show that the type of waste site may be one important explanatory factor for heterogeneity in the effect sizes observed. Considering study means instead of single estimates further supports the impression of an asymmetric plot.

Although funnel plots are an informative visual tool, their interpretation remains subjective. A regression-based formal test framework known as FAT-PET-PEESE builds on the rationale of the funnel plot, adding in the first place the standard error *SE* of the estimated effects to a simple version of eq. (5) or (6), or

$$\text{Elasticity}_i = \alpha_0 + \alpha_1 SE_i + \varepsilon_i \tag{8}$$

In this setting, the so-called Funnel Asymmetry Test (FAT) tests the hypothesis of $\alpha_1 = 0$ with a conventional t-test, assuming that, in the absence of publication bias, the effect size will be uncorrelated with its standard errors (Stanley and Doucouliagos 2012). Thus, rejecting the FAT hypothesis confirms that publication selection places a bias on the estimates in the meta-sample. Similarly, the Precision Effect Test (PET) tests for the presence of a genuine average effect size beyond publication bias ($H_0 : \alpha_0 = 0$) (Stanley 2008, 2017). In the case of a true non-zero effect confirmed by the PET, simulations have shown that the estimated average effect size $\widehat{\alpha_0}$ is often underestimated (Stanley and Doucouliagos 2012, 2017). In these cases, replacing the standard error *SE* in Eq. (8) by its square produces less biased estimates of the true underlying effect, an approach known as Precision Effect Estimate with Standard Error (PEESE) (Stanley and Doucouliagos 2014). However, if the mean effect is not significantly different from zero, the PET is shown to be the better choice. Accordingly, I estimate both alternative specifications for this meta-analysis. However, while the FAT-PET-PEESE framework is a commonly applied control for publication

bias, some simulation studies question its performance under certain conditions (Alinaghi and Reed 2018; Carter et al. 2019; Du et al. 2017). For this reason, I additionally check for the presence of publication bias by applying the publication bias control methods recently proposed by Andrews and Kasy (2019), Furukawa (2019), Ioannidis et al. (2017), Simonsohn et al. (2014) and Stanley et al. (2010).[20] For conciseness, these methods and their results are set out in detail in section B of the supplementary material.

5 Results and Discussion

The presentation of results follows the shape of the remarks on model-selection strategy set out in the previous section. The results of the tests for publication bias come first, followed by the results for the models chosen, including subsample analyses and related robustness checks. The discussion of BT errors concludes this section.

5.1 Publication Bias and Corrected Mean Effect Size

The results of the tests for publication bias are summarised in Table 3. In all cases, the weighted mean effect size is reported along with the respective coefficient controlling for publication bias, where applicable.

Regardless of the chosen method, the results show that publication bias clearly distorts the average effect size. In comparison to the unweighted average in Table 2 (0.042) and the weighted average reported in columns (1) and (4) of Table 3 (0.030 and 0.024), all methods correct the effect downwards, with estimates ranging from 0.015 to 0.029. As expected, the FAT-PET estimates (0.015 and 0.019) correct more strongly than the PEESE alternatives (0.023 and 0.029), with the PEESE estimates being less biased with regard to the significant mean effect size. In economic terms, the FAT-PET-PEESE range of estimates translates into an average increase of 1.5% to 2.9% in property values per mile of increased distance from a waste site for a house located one mile away from the waste site. For a house located 4.29 miles away from the waste site (the mean distance in this sample), the increase in value is 0.35% to 0.68%. These findings are at the lower bound of the results found in previous meta-analyses on the topic (see Sect. 2). In summary, Table 3 confirms a minor negative effect of waste sites on proximate residential property values at the aggregate level. Publication selection is present, however, resulting in an upward bias of up to 38% in this literature. This finding is corroborated by results for other publication bias control methods (see section B in the supplementary material).

Despite these findings at the aggregate level, the average effect size is based on a heterogeneous set of observations as indicated by Q-tests, funnel plots and summary statistics. I analyse the origin of this heterogeneity in the next section. As discussed previously, publication bias correction via the FAT-PET or the PEESE approach can easily be included in MRAs explaining heterogeneity. Where publication bias is clearly confirmed, *fixed effects* models should be the models of choice. On the other hand, Q-test statistics indicate a rejection of effect-size homogeneity that would appear to favour the *random effects* models.[21] In

---

[20] See also Christensen and Miguel (2018) and van Aert and van Assen (2018) for a summary of approaches addressing publication bias.

[21] The Q-test statistics are available on request.

Table 3 Results for tests of publication bias and genuine effect size

| | Meta-analytic model and publication bias control | | | Fixed effects | | |
| | Random effects | | | | | |
	WLS-RE	WLS-RE FAT-PET	WLS-RE PEESE	WLS-FE	WLS-FE FAT-PET	WLS-FE PEESE
Mean effect	0.030***	0.019***	0.029***	0.024***	0.015***	0.023***
	(0.006)	(0.007)	(0.006)	(0.005)	(0.005)	(0.005)
SE		0.382***			0.426***	
		(0.129)			(0.128)	
SE^2			0.660**			0.857**
			(0.280)			(0.350)
N	727	727	727	727	727	727
R^2	0.093	0.113	0.099	0.081	0.113	0.091

The table presents the results of WLS regressions from Eq. (8). The FAT-PET framework uses *SE*, while the PEESE correction mechanism uses SE^2. The weights are the logarithmised inverse of the squared standard error (*fixed effects*) and a composite weight also reflecting between study variance (*random effects*). Standard errors clustered at the study level are given in parentheses below the coefficients

***, ** and * denote coefficients significantly different from zero at the 1-, 5- and 10-percent levels

Table 4 Meta-regression results

Variables	Full sample			Subsamples			
	(1) WLS-RE PEESE	(2) WLS-FE PEESE	(3) G-S	(4) Mean distance	(5) Socio-economic controls	(6) Non-hazardous site	(7) On NPL
Mean effect	0.106***	0.102***	0.074***	0.098***	0.111**	−0.187	0.424**
	(0.032)	(0.031)	(0.021)	(0.034)	(0.045)	(0.144)	(0.164)
SE^2	0.847***	0.972***	0.856***	1.612**	1.030***		1.282*
	(0.279)	(0.369)	(0.271)	(0.790)	(0.179)		(0.633)
SE						0.390	
						(0.906)	
Published	0.007	0.006		0.018**	0.008	−0.031	0.097
	(0.009)	(0.006)		(0.009)	(0.010)	(0.041)	(0.070)
Year_publish	0.001*	0.001*	0.001**	0.002**	−0.004***	0.007*	0.015*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.003)	(0.007)
Site Characteristics							
Multiple sites	0.037**	0.037**	0.027***	0.028**	0.071*	−0.012	−0.002
	(0.015)	(0.014)	(0.008)	(0.014)	(0.036)	(0.047)	(0.064)
NPL (Reference Category: Not USA)							
Not on NPL	0.036	0.036		0.026	0.038		
	(0.039)	(0.033)		(0.035)	(0.038)		
On NPL	0.035	0.044		0.024	0.039		
	(0.044)	(0.036)		(0.040)	(0.040)		
Unclear	0.046	0.039		0.026	0.026		
	(0.046)	(0.039)		(0.040)	(0.041)		
Activity Status (Reference Category: Active)							
Inactive	−0.023	−0.014	−0.026	−0.019	−0.072	−0.024	−0.340***
	(0.029)	(0.025)	(0.020)	(0.028)	(0.057)	(0.046)	(0.045)
Unclear	−0.039	−0.018	−0.045***	−0.021	−0.066	−0.050	
	(0.033)	(0.019)	(0.016)	(0.031)	(0.046)	(0.092)	

Table 4 (continued)

Variables	Full sample			Subsamples			
	(1) WLS-RE PEESE	(2) WLS-FE PEESE	(3) G-S	(4) Mean distance	(5) Socio-economic controls	(6) Non-hazardous site	(7) On NPL
Employment Opportunities (Reference Category: Employment Opportunities)							
No Employment opportunities	−0.016	−0.020		−0.019	−0.023	0.120**	−0.024
	(0.024)	(0.017)		(0.025)	(0.033)	(0.047)	(0.121)
Unclear	−0.007	−0.021		−0.020	0.019		−0.075
	(0.031)	(0.022)		(0.032)	(0.039)		(0.095)
Type of Waste Site (Reference Category: Hazardous)							
Non-hazardous	−0.042	−0.041*	−0.036*	−0.040	−0.087		
	(0.026)	(0.024)	(0.022)	(0.025)	(0.055)		
Nuclear	−0.009	0.009	−0.020	−0.067	0.104		
	(0.069)	(0.062)	(0.059)	(0.099)	(0.079)		
Waste-affected element (reference category: Soil)							
Air	0.042**	0.038**	0.042**	0.022	0.011	−0.012	0.037
	(0.018)	(0.017)	(0.016)	(0.017)	(0.028)	(0.033)	(0.052)
Water	−0.007	−0.005	−0.009	−0.029*	0.005	−0.060	0.012
	(0.018)	(0.017)	(0.018)	(0.016)	(0.013)	(0.058)	(0.008)
Unclear	−0.018	−0.023	−0.016	−0.021	−0.060**	−0.048*	0.043
	(0.015)	(0.015)	(0.012)	(0.014)	(0.026)	(0.026)	(0.042)
Continent (Reference category: Europe)							
North America	−0.033	−0.031		−0.013	−0.011	0.148	
	(0.043)	(0.037)		(0.038)	(0.056)	(0.097)	
Asia	−0.015	−0.017		−0.009	0.003	0.144	
	(0.027)	(0.024)		(0.027)	(0.044)	(0.121)	
Other	0.003	0.011		−0.015	0.016	−0.027	

Table 4 (continued)

Variables	Full sample			Subsamples			
	(1) WLS-RE PEESE	(2) WLS-FE PEESE	(3) G-S	(4) Mean distance	(5) Socio-economic controls	(6) Non-hazardous site	(7) On NPL
	(0.024)	(0.020)		(0.023)	(0.038)	(0.099)	
Clean-up stage (Reference category: Recognised, no clean-up plan exists)							
Not recognised	−0.007	−0.001	−0.005	−0.020	−0.002		
	(0.014)	(0.011)	(0.012)	(0.022)	(0.013)		
Clean-up begun, but not finished	−0.007	−0.016	−0.006	−0.019	0.010		−0.012**
	(0.015)	(0.015)	(0.012)	(0.020)	(0.018)		(0.006)
Cleaned-up	−0.058**	−0.056**	−0.060**	−0.081***	−0.051*		−0.051
	(0.028)	(0.026)	(0.026)	(0.028)	(0.030)		(0.044)
Unclear	0.005	0.007	0.006	−0.009	−0.010		0.232**
	(0.018)	(0.016)	(0.016)	(0.016)	(0.026)		(0.075)
Data Characteristics							
Dist_greater_mean	−0.024*	−0.020	−0.021		−0.035***	−0.053	−0.077***
	(0.014)	(0.014)	(0.013)		(0.011)	(0.055)	(0.016)
Dist_mean				0.001			
				(0.002)			
Sample	−0.001***	−0.001***	−0.001***	−0.001***	−0.001***	0.001	−0.002**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.002)	(0.001)
Sale_ind	−0.025	−0.020		−0.021	−0.022	0.029**	−0.010
	(0.018)	(0.017)		(0.018)	(0.025)	(0.014)	(0.129)
Methodology: Econometric Specification							
Num_expl	0.000	0.000		−0.000	0.001**	−0.001	−0.002*
	(0.000)	(0.000)		(0.000)	(0.000)	(0.003)	(0.001)

Table 4 (continued)

Variables	Full sample			Subsamples			
	(1) WLS-RE PEESE	(2) WLS-FE PEESE	(3) G-S	(4) Mean distance	(5) Socio-economic controls	(6) Non-hazardous site	(7) On NPL
Oth_disamen	0.002	−0.000		−0.000	0.004	−0.108***	−0.026
	(0.011)	(0.010)		(0.012)	(0.020)	(0.033)	(0.090)
Oth_amen	−0.000	−0.003		0.004	0.006	0.021	0.029
	(0.017)	(0.018)		(0.016)	(0.028)	(0.047)	(0.065)
Access	−0.010	−0.012		−0.016	−0.002	0.064*	−0.046
	(0.012)	(0.011)		(0.012)	(0.015)	(0.035)	(0.045)
Industry	−0.025	−0.021		−0.021	−0.035	0.074	0.086
	(0.017)	(0.015)		(0.016)	(0.026)	(0.063)	(0.086)
Demoecon	−0.031**	−0.027*	−0.021*	−0.025*		−0.022	−0.017
	(0.014)	(0.015)	(0.011)	(0.014)		(0.063)	(0.018)
Time_control	0.022	0.018		0.024	−0.018	−0.010	0.007
	(0.016)	(0.015)		(0.015)	(0.026)	(0.056)	(0.049)
Direction	0.010	0.017		0.013	0.011	0.120**	0.026
	(0.017)	(0.016)		(0.015)	(0.017)	(0.055)	(0.023)
Interaction	−0.045***	−0.043***	−0.047***	−0.047***	−0.055***	−0.054	−0.016
	(0.013)	(0.012)	(0.015)	(0.013)	(0.015)	(0.057)	(0.015)
Methodology: Estimation Strategy							
Log_log	−0.012	−0.011		−0.005	−0.020	0.000	0.039
	(0.012)	(0.012)		(0.012)	(0.014)	(0.012)	(0.027)
OLS_spatial	−0.009	−0.015		−0.009	0.001	0.131	−0.084
	(0.027)	(0.025)		(0.027)	(0.030)	(0.106)	(0.089)
Observations	727	727	727	548	544	125	184
R-squared	0.261	0.232	0.246	0.343	0.277	0.610	0.613

Table 4 (continued)

Regression results are based on WLS. Standard errors clustered at the study level are given in parentheses below the coefficients.

***, ** and * denote coefficients significantly different from zero at the 1-, 5- and 10-percent levels. In columns (1) to (3), the full sample is used with different estimators and regressors, while columns (4) to (7) depict subsample results

such a case, the literature provides no clear guidance on which estimator to prefer and the choice of the appropriate model follows the decision rules described in Sect. 4.1.

5.2 Heterogeneity of Effect Size

The meta-regression results are reported in Table 4. Column (1) shows the baseline WLS-RE PEESE model. Selection of this specification follows the results of a Q-test rejecting the null hypothesis of no heterogeneity at the estimate level and a subsequent BPLM test lending no support for the hypothesis of additional study-level heterogeneity.[22] In summary, the moderators included are sufficient to explain study-level heterogeneity so that there is no need to rely on panel-econometric models to reflect study-level effects.[23] For this baseline specification I consider all moderators with non-missing observations defined in Table 1 so as to conserve sample size. The PEESE publication bias control is included as this is the preferred choice with a significant mean effect size resulting in a smaller bias, as set out below. Results including FAT-PET publication bias control are reported in Table A8 in the supplementary material.

I have checked the baseline model for normality of residuals, outliers, persistent heteroscedasticity and multicollinearity. The findings support inference validity; see Figures A5 and A6 in the supplementary material for details on non-normality and outliers.[24] However, persistent heteroscedasticity was identified despite reliance on WLS. Consequently, cluster-robust standard errors are used for all regressions. Multicollinearity was only a minor concern that did not affect the regression results in any relevant way.[25]

I explore the robustness of the results from a variety of perspectives. Column (2) represents the WLS-FE model, reflecting the discussion in the previous section on the preferred estimator in the presence of publication bias. The results for a reduced model are reported in column (3), following a general-to-specific (G-S) modelling approach recommended by Stanley and Doucouliagos (2012). This approach involves a stepwise removal of the least significant variable until only variables with a p-value less than 0.2 remain. Additional regressions based on trimmed datasets and reweighted observations confirm the robustness of the results. They are shown in Table A8 in the supplementary material.

The majority of moderators are binary or categorical variables. Accordingly, their corresponding coefficients can be interpreted *ceteris paribus* as the expected change in mean effect size caused by a departure from the benchmark scenario. The benchmark scenario for the categorical moderators is the omitted category indicated in parentheses. For all

---

[22] Here, the BPLM test cannot be performed in its standard version. However, this can easily be resolved by slightly altering the approach; see the detailed explanations in the Do-files made available in the data repository linked to in the electronic version.

[23] Detailed test results are summarised in Table A7.

[24] These exemplary diagnostic plots for the WLS-RE PEESE model detect six potentially outlying observations. I reassessed them for coding errors or any particularities justifying their exclusion. No obvious pattern was detected. As the regression results are robust to the exclusion of the outlying observations, they are not discarded. See Table A8 in the supplementary material for details. Diagnostic plots for the other regression models based on the entire sample come up with almost identical results and are available on request.

[25] The mean variance inflation factor (VIF) is 3.7 with only a few levels of some categorical regressors indicating multicollinearity. These are to be expected as *e.g.* the status on the NPL is, by design, correlated with the status *North_America*. Reassuringly, the reduced model in column (3) and the clear majority of robustness regressions confirm the same set of significant moderators, so that multicollinearity does not affect the results in any relevant way.

binary regressors, the benchmark is the zero case. The remaining continuous variables are centred so that their coefficients can be interpreted as the effect of deviations from the mean.[26] Hence, the constant can be interpreted as the mean effect size for a reference study indicated by the benchmarks. Table 4 shows that the results are robust across specifications based on all observations. With only small quantitative differences for most coefficients, the following description focuses on the results of the WLS-RE model and only has recourse to the WLS-FE and G-S alternatives in the case of pronounced disagreement.

Most notably, publication bias is confirmed throughout the models. This squares with the results from Tables 3 and B1, finding evidence of a highly significant upward bias in the literature. The corrected mean effect sizes in columns (1) to (3) range from about 0.074 to 0.106 in magnitude and are significant in all cases. In comparison to mean effect sizes displayed in Table 2 and Table 3, mean effect sizes in the comprehensive meta-regressions are two to three times larger. This can be explained by the different cases reflected in the respective mean effect sizes.[27] The R^2 shows little variance, ranging between 0.232 and 0.261. Turning to the explanatory variables, published and unpublished studies do not seem to differ in effect size magnitude when publication bias is controlled for. More recent publications tend to report greater effect sizes.

5.2.1 Site Characteristics

As expected, studies with multiple sites in the proximity of residential properties report higher effect sizes on average. In other words, multiple waste sites affect residential property values more adversely than single sites. By contrast, there seem to be no significant differences in terms of effect size for waste sites with differing employment opportunities or status on the NPL. According to the reduced G–S model, sites with unclear activity status seem to have a less value-depressing impact than active sites. Surprisingly, the type of waste does not seem to influence effect size in any consistent way. Only in the WLS-FE and G–S model are the effects of non-hazardous waste sites significantly smaller than their hazardous counterparts (and only at the 10% significance level). I explore this result in more detail in the subsample regressions discussed below. Further, there seem to be distinctive differences in effect size depending on the element affected by waste. Waste sites emitting airborne pollutants clearly reduce residential property values more significantly than waste sites where the polluted element is unclear, whereas soil-polluting or water-polluting waste sites do not significantly differ from the latter in terms of effect size. The continent on which a waste site is located does not seem to be of relevance. In neither case is there any significant difference from a European waste site.[28] The clean-up stage seems to have a distinct impact on property values. Recalling the reference case of a waste site with recently discovered pollution, it transpires that, during the remediation phase, there is no significant recovery in residential property values. However, this recovery effect manifests

---

[26] This excludes, of course, the publication bias control variables *SE* and *SE²*.

[27] The reference study in the meta-regressions displayed in Table 4 is, *e.g.,* unpublished, does not include socio-demographic controls in its regression(s) and models the effect of a European waste site, as indicated by the definition of the moderators. As the majority of the studies in the meta-dataset are published, include socio-demographic controls and focus on North American waste sites, the mean effect sizes in Table 2 and Table 3 reflect a different reference case.

[28] One reviewer suggested replacing the continent dummies with a continuous measure like the HDI or GDP per capita. Both replacements are tested and reported in Table A9. The results are similar to the baseline specification.

itself once the clean-up has been completed. This finding, intuitive as it is, contrasts with the results of Braden et al. (2011), who report insignificant effects from clean-up activities.

5.2.2 Data Characteristics

The data characteristics reveal that on average studies working with larger samples tend to report smaller estimates. By contrast, studies based on sales data collected at individual house level do not differ significantly from studies using assessed values or aggregated data. Apparently, studies with greater mean distances of the sampled houses from the respective waste site do not consistently report significantly smaller effect sizes than studies with smaller mean distances. Only in the WLS-RE model, the coefficient is significant, albeit weakly. This finding is seemingly at odds with the basic hedonic hypothesis that an environmental disamenity will be considered less serious the further away it is from the property in question. However, the effect may be obscured by the choice of the dummy variable *dist_greater_mean* as it is only a crude measure of the distance-decay effect. In addition, reported effect sizes may be biased if primary studies fail to control for important confounding factors. Finally, if on average non-hazardous waste sites are not perceived as disamenities, this may also mask the distance-decay effect. I investigate these eventualities in the subsample analyses.

5.2.3 Methodology

Whereas the mere number of explanatory variables does not seem to be an important factor, some moderators reflecting the comprehensiveness or quality of the econometric specification do significantly affect the reported effect size.[29] First, not controlling for socio-demographic factors, such as the crime rate or median household income in a census tract, results in significantly higher estimates of waste-site effects. This shows that the omission of important control variables in the primary literature potentially leads to biased estimates of the effect of waste sites on residential property values. However, controlling for other amenities or non-waste disamenities in the vicinity does not on average appear to influence the reported effect size.

Moreover, studies that use a distance-interaction term report considerably smaller average effect sizes. Evidently, the interaction term takes up some part of the effect size that would otherwise be reflected in the single term. By contrast, controlling for time-fixed effects via dummy variables or using price inflation adjustments does not seem to influence the reported effect size. Lastly, the estimation strategy is not a significant factor explaining variance. More precisely, neither the functional form nor the chosen estimator seem to influence the reported effect size in any given case.

With very similar results across specifications and estimators, the subsample analyses rely on the WLS-RE PEESE baseline model. Based on the results from Table 3 and Table 4, this is justifiable as the mean effect size consistently differs significantly from zero.

---

[29] In a robustness check reported in Table A9, the number of total explanatory variables is replaced by the number of significant variables, which is also insignificant. The overall results remain unchanged except for an increased R^2. This effect, however, is not due to the exchanged variable but rather related to the effect of a smaller, more consistent subsample.

5.3 Subsample Analysis

Several moderators are not included in the baseline regression shown in column (1) due to missing observations or because these moderators only serve as replacements for explanatory variables already included. In a first step, I add these moderators to the meta-regression separately, which naturally leads to a reduced sample size in each case. Moreover, the results from the baseline model identify sub-categories of observations that warrant further investigation. Accordingly, I use different subsample regressions to provide a more nuanced picture of the robustness of results. In most cases, the subsample analyses reveal only marginal differences from the baseline model. Reassuringly, however, expected differences manifest themselves as well. The results for four subsamples are summarised in columns (4) to (8) in Table 4. For other subsample regressions, see Table A9 in the supplementary material. Results are generally robust, so again I only discuss notable differences from the baseline model.

As expected, replacing *Dist_greater_mean* with *Dist_mean* in column (4) makes very little difference to the overall results. Notably, the coefficient on *Dist_mean* is insignificant, reproducing the counterintuitive finding of Braden et al. (2011). Hence, a continuous distance-decay effect cannot be confirmed for the full sample. When we turn to the subsamples in columns (5) to (7), the differences from the baseline model are more clearcut. Controlling for socio-economic factors (column (5)) seems to be an important quality dimension in the primary study. While the overall results remain similar to the baseline regression, the distance-decay effect is now confirmed. This finding manifests the impression that the design of the primary study may be one factor concealing this pattern. This interpretation is supported by an additional subsample regression that omits effect sizes from interactions (shown in Table A9).

For the subsample of non-hazardous waste-site observations in column (6), the mean effect size turns insignificant. On average, non-hazardous waste sites are apparently not value-depressing. This finding supports the hypothesis that waste-site effects on property values differ by waste categories that can be controlled for in MRAs. Similarly, no publication bias can be confirmed. Neither the dummy variable *Publish* nor the publication bias control variable *SE* significantly affect the average effect size. It appears that there are no prior expectations of the sign or significance of the effect size for this type of waste site. Moreover, in contrast to the baseline model, other disamenities in the vicinity significantly reduce the effect size of non-hazardous waste sites. In addition, a non-hazardous waste site offering no employment is clearly more value-depressing than otherwise. The R^2 increases substantially compared to the baseline scenario (0.610 compared to 0.261).

Surprisingly, status on the NPL seemed to be of no relevance for the effect size in any of the previous regressions. Hence, in column (7) I explore the subsample of observations from waste sites on the NPL more closely. The mean effect size is approximately fourfold in magnitude, confirming the expected negative effects of highly contaminated waste sites. Moreover, the R^2 increases substantially compared to the baseline scenario (0.613 compared to 0.261). The clean-up stage continues to be of importance. More precisely, for this subset of highly contaminated sites, an unclear clean-up stage increases effect size over and against a site with recently discovered contamination. In addition, the start of clean-up activities seems to be an important step towards remediation. As expected, the

distance-decay effect can be confirmed.[30] Apparently, the type and level of contamination are determinants in detecting this pattern (see also the subsample regression on hazardous waste sites in Table A9).

5.4 Benefit Transfer

As noted previously, one of the potential merits of MRA is the use of the resulting coefficients for BT applications. The usefulness for BT applications, however, depends on the magnitude of the inherent transfer error. Table 5 shows the transfer errors to illustrate the predictive power of this MRA. I calculate the transfer errors following common practice, using the Absolute Percentage Error (APE) as metric (Nelson 2015). The APE is shown for the baseline specification that includes all observations as well as for the subsamples of non-hazardous waste sites and waste sites on the NPL, respectively.[31] Moreover, the APE is calculated using not only meta-functional transfer based on the respective results shown in Table 4, but also mean-value transfer based on a univariate WLS analogue of Eq. (6). Following Brander et al. (2006) and Chaikumbung et al. (2016),[32] I calculate the transfer errors using n-1 out-of-sample regressions, *i.e.*, omitting one observation at a time, re-estimating the model and calculating the estimated BT for the observation omitted.

The general impression gained from the results on transfer errors is in line with previous discussions in the literature. First, meta-functional transfer outperforms simple value transfer if the underlying sample is heterogeneous. If the sample consists of a more homogenous set of sites, *e.g.*, waste sites on the NPL, simple value transfer results in smaller transfer errors. Second, regardless of the type of transfer, the APE is smaller if the pooled sites have a higher degree of commodity consistency. Third, the Mean APE is considerably larger than the Median APE, with BT errors ranging between 133% and 684% for the former, compared to values of 76% to 90% for the latter. A small number of outlying observations drives the mean APE upwards.[33] This pattern is also apparent in Figure A7 in the supplementary material, showing the distribution of BT errors.

The results are in line with expectations, but the order of magnitude of the BT errors warrants further investigation. In general, moderate to high transfer errors are common in the valuation literature. Rosenberger (2015) reviews the transfer errors for 38 valuation studies, reporting an average of 65% (36%) for the Mean (Median) APE based on function transfer and a corresponding average of 140% (45%) for value-transfer errors. However, though the BT errors shown in Table 5 are greater than this average in most cases, this is not necessarily surprising as the complex model was designed to identify sources of heterogeneity rather than to calculate transfer errors (Boyle and Wooldridge 2018; Nelson 2015). Still, while the acceptable level of transfer error is context-dependent (Rosenberger

---

[30] For this subsample regression, the moderator *Dist_mean* was also significant (not shown).

[31] In all cases, one observation with an extreme APE (770,261%) is omitted from the calculations and from the presentation in Table 5. Omitting this clearly outlying observation reduces *e.g.* the Mean APE from 1,602% to 543% for the full sample, while the Median APE barely changes. Renewed perusal of the corresponding paper (Schmalensee 1975) revealed no particularities.

[32] I calculated in-sample transfer error as well (not shown). They follow the same pattern as the out-of-sample predictions depicted in Table 5, albeit with smaller transfer errors throughout, as expected.

[33] Omitting 14 observations from one study (Walsh and Mui 2017) that represent the vast majority of extreme APE reduces the Mean APE from 543% to 213%. Similarly, for the subset of observations on the NPL, four observations are clearly outlying. Omitting these observations reduces the corresponding Mean APE from 359% to 125%.

Table 5 Benefit transfer errors

	Mean APE (SD)	Median APE	Range of APE
Meta-Functional Transfer			
Full sample	543.41 (3228.37)	87.55	0.63–49,363
Non-hazardous waste sites	463.90 (1708.55)	84.52	0.93–13,253
Waste sites on the NPL	358.71 (2224.52)	90.45	0.19–29,098
Mean Value Transfer			
Full sample	683.56 (4635.68)	85.92	0.03–72,724
Non-hazardous waste sites	601.33 (2234.48)	79.72	2.78–15,146
Waste sites on the NPL	133.08 (333.77)	75.92	1.30–3575

The meta-functional transfer calculations are based on the regressions in columns (1), (6) and (7) from Table 4. The mean value transfer is based on the univariate WLS-RE analogue of Eq. (6)

2015; Brander et al. 2006), the predicted levels of transfer error will in most cases prohibit accurate policy applications. Apparently, waste-site effects on property values are difficult to predict reliably.[34] Nevertheless, the BT estimates may still be informative for broad cost–benefit analysis at policy sites where primary studies are not feasible.

6 Conclusion

This meta-study confirms the existence of adverse price effects on residential property values caused by the proximity to waste sites at the aggregate level. Correcting for publication bias has a sizeable impact, reducing the average effect size by up to 38%. The corrected average effect size translates into a 1.5% to 2.9% property value increase per mile of increased distance from a waste site for a house at a one-mile distance. These estimates are situated in the lower range of values produced by the previous literature. The results are generally robust across justifiable estimators, weighting schemes and the replacement of moderators. This need not necessarily hold in other circumstances, and future researchers would do well to adhere to the structured decision pathways that already exist (Feld and Heckemeyer 2011; Stanley and Doucouliagos 2012) in choosing the appropriate model for their respective meta-dataset.

Various site and data characteristics as well as the respective econometric specification are significant factors explaining the variation in the empirical findings. Notably, the distance decay of the waste-site effect is partly confirmed, *e.g.*, for hazardous waste sites in general and waste sites on the NPL in particular. In addition, cleaning up a waste site is beneficial for residential property values, possibly restoring value formerly forfeited. This finding contrasts with previous insignificant findings by Braden et al. (2011), possibly due to smaller sample size or the absence of subsample analysis for hazardous waste sites. The subsample analyses revealed distinct differences for severely contaminated sites on the NPL and non-hazardous waste sites. As non-hazardous waste sites do not reduce property values on average, they are not considered a disamenity in these average cases. By contrast,

---

[34] This interpretation is supported by additional calculations of BT errors for even more consistent subsamples like landfills or brownfields. The transfer errors were in the same range.

severely contaminated waste sites on the NPL clearly reduce residential property values on average, with an estimated mean effect size of 42.2%. Future MRAs in this context may want to concentrate separately on these types of waste site, as this might be a way of identifying more waste-type-specific moderators. This might also increase forecast accuracy for BT applications. The BT errors in this MRA are consistent with general findings in the literature, showing that meta-functional transfer performs better than value transfer if sites are heterogeneous. The practical applicability of the BT estimates, however, is limited due to comparatively high transfer errors. Hence, more specialised MRAs focusing on waste sites with a higher degree of similarity and corresponding waste-type-specific moderators are needed to forecast estimates more reliably.

There are at least two avenues for future research to explore. First, new moderators need to be identified that can help shed light on the remaining unexplained variance. Second, there is still no unambiguous picture of the presumed distance-decay effect. With occasional insignificance of the moderator controlling for the mean distance between houses and waste sites in at least two MRAs, this remains a partly unresolved issue. Clearly, it would be of great interest to find average distance cut-off points beyond which a waste site is no longer perceived as a disamenity.

Compliance with Ethical Standards

Conflict of interest The author declares that he has no conflict of interest.

References

Affuso E, Hite D, Ho C-S, Pariso CVd (2010) The impact of hazardous waste on property values: the effect of lead pollution. Urbani Izziv 21:117–126

Alinaghi N, Reed WR (2018) Meta-analysis and publication bias: how well does the FAT-PET-PEESE procedure work? Res Synth Methods 9:285–311

Andrews I, Kasy M (2019) Identification of and correction for publication bias. Am Econ Rev 109:2766–2794

Bateman IJ, Brouwer R, Ferrini S, Schaafsma M, Barton DN, Dubgaard A, Hasler B, Hime S, Liekens I, Navrud S, de Nocker L, Ščeponavičiūtė R, 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. Environ Resour Econ 50:365–387

Beltrán A, Maddison D, Elliott RJR (2018) Is flood risk capitalised into property values? Ecol Econ 146:668–685

Bergstrom JC, Taylor LO (2006) Using meta-analysis for benefits transfer: theory and practice. Ecol Econ 60:351–360

Bilbao-Terol C (2009) Impacts of an iron and steel plant on residential property values. Eur Plan Stud 17:1421–1436

Borenstein M, Hedges L, Higgins J, Rothstein H (2011) Introduction to meta-analysis. Wiley, Chichester, UK

Boyle M, Kiel K (2001) A survey of house price hedonic studies of the impact of environmental externalities. J Real Estate Lit 9:117–144

Boyle KJ, Wooldridge JM (2018) Understanding error structures and exploiting panel data in meta-analytic benefit transfers. Environ Resour Econ 69:609–635

Boyle KJ, Parmeter CF, Boehlert BB, Paterson RW (2013) Due diligence in meta-analyses to support benefit transfers. Environ Resour Econ 55:357–386

Braden JB, Feng X, Won D (2011) Waste sites and property values: a meta-analysis. Environ Resour Econ 50:175–201

Braden JB, Feng X, Freitas L, Won D (2010) Meta-functional transfer of hedonic property values: Application to Great Lakes Areas of Concern. Agricul Resour Econ Rev 39:101–113

Brander LM, Florax RJGM, Vermaat JE (2006) The empirics of wetland valuation: a comprehensive summary and a meta-analysis of the literature. Environ Resour Econ 33:223–250

Brinkley C, Leach A (2019) Energy next door: a meta-analysis of energy infrastructure impact on housing value. Energy Res Soc Sci 50:51–65

Brouwer R, Sheremet O (2017) The economic value of river restoration. Wat Resour Econ 17:1–8

Card N, Little T (2016) Applied meta-analysis for social science research. Methodology in the social sciences. The Guilford Press, New York

Carter EC, Schönbrodt FD, Gervais WM, Hilgard J (2019) Correcting for bias in psychology: a comparison of meta-analytic methods. Adv Meth Pract Psychol Sci 2:115–144

Chaikumbung M, Doucouliagos H, Scarborough H (2016) The economic value of wetlands in developing countries: a meta-regression analysis. Ecol Econ 124:164–174

Chen WY, Li X, Hua J (2019) Environmental amenities of urban rivers and residential property values: a global meta-analysis. Sci Tot Environ 693:133628

Chèze B (2007) Une méta-analyse des études d'évaluation monétaire par la méthode des prix hédoniques du coût externe des installations de traitement des déchets (No. 2007–23). University of Paris Nanterre, EconomiX

Christensen G, Miguel E (2018) Transparency, reproducibility, and the credibility of economics research. J Econ Lit 56:920–980

Debrezion G, Pels E, Rietveld P (2007) The impact of railway stations on residential and commercial property value: A meta-analysis. J Real Estate Financ 35:161–180

Du H, Liu F, Wang L (2017) A Bayesian, "fill-in" method for correcting for publication bias in meta-analysis. Psychol Methods 22:799–817

Du Preez M, Balcilar M, Razak A, Koch SF, Gupta R (2016) House values and proximity to a landfill in South Africa. J Real Estate Lit 24:133–149

Environmental Protection Agency (2019) Basic NPL Information. https://www.epa.gov/superfund/basic-npl-information. Accessed 29 Jan 2018

Eshet T, Baron MG, Shechter M (2007a) Exploring benefit transfer: disamenities of waste transfer stations. Environ Resour Econ 37:521–547

Eshet T, Baron MG, Shechter M, Ayalon O (2007b) Measuring externalities of waste transfer stations in Israel using hedonic pricing. Waste Manage 27:614–625

European Commission (1998) Undesirable facilities and property values: a summary of empirical studies. Ecol Econ 24:1–14

Farber S (1998) Undesirable facilities and property values: a summary of empirical studies. Ecol Econ 24:1–14

Feld LP, Heckemeyer JH (2011) FDI and taxation: a meta-study. J Econ Surv 25:233–272

Furukawa C (2019) Publication bias under aggregation frictions: theory, evidence, and a new correction method

Giusti L (2009) A review of waste management practices and their impact on human health. Waste Management 29:2227–2239

Gunby P, Jin Y, Reed WR (2017) Did FDI really cause Chinese economic growth? A meta-analysis. World Dev 90:242–255

Ioannidis JPA, Stanley TD, Doucouliagos H (2017) The power of bias in economics research. Econ J 127:F236–F265

Jackson T (2001) The effects of environmental contamination on real estate: a literature review. J Real Estate Lit 9:91–116

Johnston RJ, Rolfe J, Rosenberger RS, Brouwer R (2015) Introduction to benefit transfer methods. In: Johnston RJ, Rolfe J, Rosenberger RS, Brouwer R (eds) benefit transfer of environmental and resource values: a guide for researchers and practitioners. Springer, Netherlands, Dordrecht, pp 19–59

Johnston RJ, Besedin EY, Stapler RW (2017) Enhanced geospatial validity for meta-analysis and environmental benefit transfer: An application to water quality improvements. Environ Resour Econ 68:343–375

Johnston RJ, Rolfe J, Zawojska E (2018) Benefit transfer of environmental and resource values: progress, prospects and challenges. IRERE 12:177–266

Kaza S, Yao L, Bhada-Tata P, van Woerden F (2018) What a waste 2.0: a global snapshot of solid waste management to 2050. Urban development series. World Bank Publications, Washington

Kiel KA, Williams M (2007) The impact of superfund sites on local property values: are all sites the same? J Urban Econ 61:170–192

Klemick H, Griffiths C, Guignet D, Walsh P (2018) Improving water quality in an iconic estuary: an internal meta-analysis of property value impacts around the Chesapeake Bay. Environ Resour Econ 69:265–292

Li C-N, Lo C-W, Su W-C, Lai T-Y (2015) A study on soil and groundwater pollution remediation of the surrounding real estate prices and tax revenue impact. Sustainability 7:1–13

Lindhjem H, Navrud S (2008) How reliable are meta-analyses for international benefit transfers? Ecol Econ 66:425–435

Lipscomb C, Mooney A, Kilpatrick J (2013) Do survey results systematically differ from hedonic regression results? Evidence from a residential property meta-analysis. J Real Estate Lit 21:233–253

Neelawala P, Wilson C, Athukorala W (2013) The impact of mining and smelting activities on property values: a study of Mount Isa city, Queensland, Australia. Aust J Agric Resour Econ 57:60–78

Nelson JP (2004) Meta-analysis of airport noise and hedonic property values: problems and prospects. Journal of Transp Econ Policy 38:1–28

Nelson JP (2015) Meta-analysis: Statistical methods. In: Johnston RJ, Rolfe J, Rosenberger RS, Brouwer R (eds) Benefit transfer of environmental and resource values: a guide for researchers and practitioners. Springer, Netherlands, Dordrecht, pp 329–356

Nelson JP, Kennedy PE (2009) The use (and abuse) of meta-analysis in environmental and natural resource economics: an assessment. Environ Resour Econ 42:345–377

Penn J, Hu W (2019) Cheap talk efficacy under potential and actual hypothetical bias: a meta-analysis. J Environ Econ Manag 96:22–35

Phaneuf D, Requate T (2017) A course in environmental economics: theory, policy, and practice. Cambridge University Press, New York, NY

Poor PJ, Pessagno KL, Paul RW (2007) Exploring the hedonic value of ambient water quality: a local watershed-based study. Ecol Econ 60:797–806

Ready R (2010) Do landfills always depress nearby property values? J Real Estate Res 32:321–339

Reichert AK, Small M, Mohanty S (1992) The impact of landfills on residential property values. J Real Estate Res 7:297–314

Ringquist E (2013) Meta-analysis for public management and policy. Wiley, San Francisco, CA

Rosen S (1974) Hedonic prices and implicit markets: product differentiation in pure competition. J Polit Econ 82:34–55

Rosenberger RS (2015) Benefit transfer validity and reliability. In: Johnston RJ, Rolfe J, Rosenberger RS, Brouwer R (eds) Benefit transfer of environmental and resource values: a guide for researchers and practitioners. Springer, Netherlands, Dordrecht, pp 307–326

Saginor J, Simons R, Throupe R (2011) A meta-analysis of the effect of environmental contamination on non-residential real estate values. J Prop Invest Finance 29:460–478

Simons R, Saginor J (2006) A meta-analysis of the effect of environmental contamination and positive amenities on residential real estate values. J Real Estate Res 28:71–104

Simonsohn U, Nelson LD, Simmons JP (2014) P-curve: a key to the file-drawer. J Exp Psychol Gen 143:534–547

Smith VK, Pattanayak SK (2002) Is meta-analysis a Noah's Ark for non-market valuation? Environ Resour Econ 22:271–296

Stanley TD (2001) Wheat from chaff: meta-analysis as quantitative literature review. J Econ Perspect 15:131–150

Stanley TD, Doucouliagos H (2012) Meta-regression analysis in economics and business. Routledge, New York, NY

Stanley TD, Doucouliagos H (2014) Meta-regression approximations to reduce publication selection bias. Res Synth Methods 5:60–78

Stanley TD, Doucouliagos H (2017) Neither fixed nor random: weighted least squares meta-regression. Res Synth Methods 8:19–42

Stanley TD, Doucouliagos H, Giles M, Heckemeyer JH, Johnston RJ, Laroche P, Nelson JP, Paldam M, Poot J, Pugh G, Rosenberger RS, Rost K (2013) Meta-analysis of economics research reporting guidelines. J Econ Surv 27:390–394

Stanley TD, Jarrell SB, Doucouliagos H (2010) Could it be better to discard 90% of the data?: a statistical paradox. Amer Stat 64:70–77

Van Aert RCM, van Assen M (2018) Correcting for publication bias in a meta-analysis with the p-uniform* method. https://osfio/preprints/bitss/zqjr92018. Accessed from 2018

Vedogbeton H, Johnston RJ (2020) Commodity consistent meta-analysis of wetland values: an iillustration for coastal marsh habitat. Environ Resour Econ 75:835–865

Walsh P, Mui P (2017) Contaminated sites and information in hedonic models: an analysis of a NJ property disclosure law. Resour Energy Econ 50:1–14

Walton H, Boyd R, Taylor T, Markandya A (2006) Explaining variation in amenity costs of landfill: meta-analysis and benefit transfer. Paper presented at the 3rd world congress of environmental and resource economists, Kyoto, July 2006

Wooldridge J (2010) Econometric analysis of cross section and panel data. MIT Press, Cambridge, MA

Zegarac M, Muir T (1998) The effects of RAP related restoration and parkland development on residential property values: a hamilton harbour case study. Environment Canada, Ontario Region

Zeiss C (1998) Cause and effect patterns of noxious facility impacts on property values. J Environ Syst 26:111–136