Measuring the Income Elasticity of Water Demand: The Importance of Publication and Endogeneity Biases*

Tomas Havranek[a,b], Zuzana Irsova[b], and Tomas Vlach[b]

[a]Czech National Bank
[b]Charles University, Prague

August 21, 2017

Abstract

We present the first study that examines the effects of publication selection in the literature estimating the income elasticity of water demand. Paradoxically, more affected by publication selection are the otherwise preferable estimates that control for endogeneity. Because such estimates tend to be smaller and less precise, they are often statistically insignificant, which leads to more intense specification searching and bias. Attempting to correct simultaneously for publication and endogeneity biases, we find that the mean underlying elasticity is approximately 0.15 or less. The result is robust to controlling for more than 30 other characteristics of the estimates and using Bayesian model averaging to account for model uncertainty. The differences in the reported estimates are systematically driven by differences in the tariff structure, regional coverage, data granularity, and control for temperature in the demand equation.

Keywords: Water demand, income elasticity, meta-analysis, publication bias, Bayesian model averaging

JEL Codes: C83, Q25

*Corresponding author: Zuzana Irsova, zuzana.irsova@ies-prague.org. Data and code are available in an online appendix at meta-analysis.cz/water. Havranek acknowledges support from Czech Science Foundation grant #15-02411S; Irsova acknowledges support from Czech Science Foundation grant #16-00027S. This project has also received funding from Charles University project "Economics of Energy and Environmental Policy" (PRIMUS/17/HUM/16), Czech Science Foundation grant #18-02513S, and the European Union's Horizon 2020 Research and Innovation Staff Exchange program under the Marie Sklodowska-Curie grant agreement #681228. We thank Maamar Sebri for providing us with the list of studies used in his meta-analysis and are grateful to two anonymous referees of Land Economics for their useful comments. The views expressed here are ours and not necessarily those of the Czech National Bank.

1 Introduction

The growing scarcity of drinking water represents a major global risk (WEF, 2015). To understand how the consumption of water will evolve when developing countries get richer, we need reliable estimates of the income elasticity of water demand. The parameter is also used by policy makers to design efficient and equitable environmental water policies. Researchers have long sought to pin down this crucial parameter but have yet to reach consensus. The two previous quantitative surveys conducted on this topic, Dalhuisen *et al.* (2003) and Sebri (2014), put the representative estimate in the literature between 0.2 and 0.4 and focus on the drivers of heterogeneity in the reported income elasticities. Neither of these studies, however, corrects the literature for publication bias, and neither accounts for model uncertainty when explaining the heterogeneity behind the estimates. In this paper we collect 307 estimates of the income elasticity of water demand and analyze the variation behind these estimates, paying special attention to publication bias, endogeneity bias, and model uncertainty.

Publication bias arises from the tendency of researchers, editors, and referees to publish results that are either significant or have the desired sign. In theory, water is a necessity with no obvious substitutes; therefore, common sense dictates that its income elasticity should be positive and statistically significant. But if the underlying elasticity that we try to unearth is sufficiently small and our data and methods sufficiently imprecise, we should get negative or statistically insignificant estimates from time to time. If such estimates are underreported, publication bias arises. In a related study on the price elasticity of water demand, Stanley (2005) finds that publication bias exaggerates the estimates *fourfold*. The studies estimating the price elasticity of water demand typically also estimate the income elasticity, often in the same equation. This demonstrates the importance of accounting for publication selection.

The endogeneity problem in water demand equations (price is not exogenous to quantity) is well documented and has been explored by previous meta-analyses. Here we offer a twist to the typical story that estimates of the income elasticity taken from models accounting for price endogeneity are always preferable. This statement holds when no publication selection exists. But if publication selection constitutes a problem, as we show is the case, estimates based on instrumental variables give rise to more publication selection because they are typically (though not always) less precise than OLS estimates and, in this particular case, also smaller.

Researchers seeking to control for endogeneity while simultaneously providing estimates that are publishable (intuitive and statistically significant) are sometimes forced to pursue a lengthy search for the desired specification with a point estimate that is large enough to offset the standard error. It follows that OLS estimates are exaggerated by endogeneity bias, while IV estimates are exaggerated by publication bias, and the simple mean reported elasticities might not vary substantially between these two approaches.

Indeed, our results suggest that the income elasticity of water demand is, on average, biased upwards due to publication bias and that the extent of bias is linked to the treatment of price endogeneity. Publication bias is absent from estimates produced by methods ignoring endogeneity (such as OLS). By contrast, while methods controlling for endogeneity (such as IV) report estimates that are hoped to be corrected for endogeneity bias, these estimates are correlated with their standard errors and collectively suffer from publication bias. As a result, although researchers address endogeneity bias at the level of individual studies, the resulting publication bias means that the mean reported estimate is not closer to the underlying value of the income elasticity. This interplay between the two biases is too complex for any narrative survey to decipher, and the use of meta-analysis is therefore crucial. The two biases cause the reported estimates to be similar for IV and OLS methods, which has led previous meta-analyses to conclude that trying to correct for price endogeneity, while theoretically laudable, has little practical effect on the income elasticity. We argue otherwise.

Furthermore, we collect 38 method and data characteristics that should help us explain the differences among the estimated elasticities. The large number of characteristics, however, means that we face model uncertainty, so we depart from the frequentist methods of the previous meta-analyses and instead apply Bayesian model averaging. Bayesian model averaging runs millions of regressions that include the possible subsets of the explanatory variables. Consequently, it constructs a weighted average over these regressions, where each weight is approximately proportional to the goodness of fit of the respective regression. The results of Bayesian model averaging enable us to construct a "best-practice" estimate in the literature conditional on numerous data and method choices, which is another value added of meta-analysis. It follows that the income elasticity of water demand is likely 0.15 or even less, smaller than usually perceived, and in any case the literature is inconsistent with values of the elasticity over 0.5.

The remainder of the paper is organized as follows. Section 2 describes the data collection approach and the basic properties of the data set. Section 3 tests for the presence of publication selection bias and explores its relationship with endogeneity bias. Section 4 investigates the data, method, and publication heterogeneity in the estimated income elasticities and constructs best-practice estimates for different pricing schemes. Section 5 presents a battery of robustness checks. Section 6 concludes the paper. An online appendix, available at meta-analysis.cz/water, provides the data and code that allow other researchers to replicate our results.

2 The Data Set

To estimate the income elasticity of water demand, researchers usually employ a variant of the following model:

$$\ln Consumption_{it} = \alpha + PED \cdot \ln Price_{it} + YED \cdot \ln Income_{it} + Controls_{ijt} + \epsilon_{it}, \qquad (1)$$

where $Consumption_{it}$ denotes the water consumption of household i in period t, *Price* denotes the price of water, and *Income* denotes household income. The vector $Controls_{ijt}$ represents a set of explanatory variables j, such as household characteristics (the number of household members, distinction between primary and secondary residences, garden size, and the number of bathrooms) or climate variables (temperature, rainfall, and evaporation). The coefficient *PED* is the price elasticity of water demand; ϵ is the error term. The coefficient *YED* denotes the income elasticity, the effect in question of this meta-analysis, and captures by how many percent the demand for water changes if the household's income increases by one percent. As in any other meta-analysis, we begin by collecting the reported estimates from the empirical literature. We exploit previously published meta-analyses by Dalhuisen *et al.* (2003) and Sebri (2014) and extend the data sample by searching the Google Scholar database; the search query is available online at meta-analysis.cz/water. We add the last study on March 6, 2016.

To be included in the meta-analysis, the studies must conform to three criteria: 1) the study must estimate a water demand equation and report an empirical estimate of *YED*; 2) the study must estimate the log-log functional form of a demand equation, as in (1), to display a constant *YED*; and 3) the study must report a measure of uncertainty around the estimate, typically the standard error. Several studies do not conform to these criteria. For example,

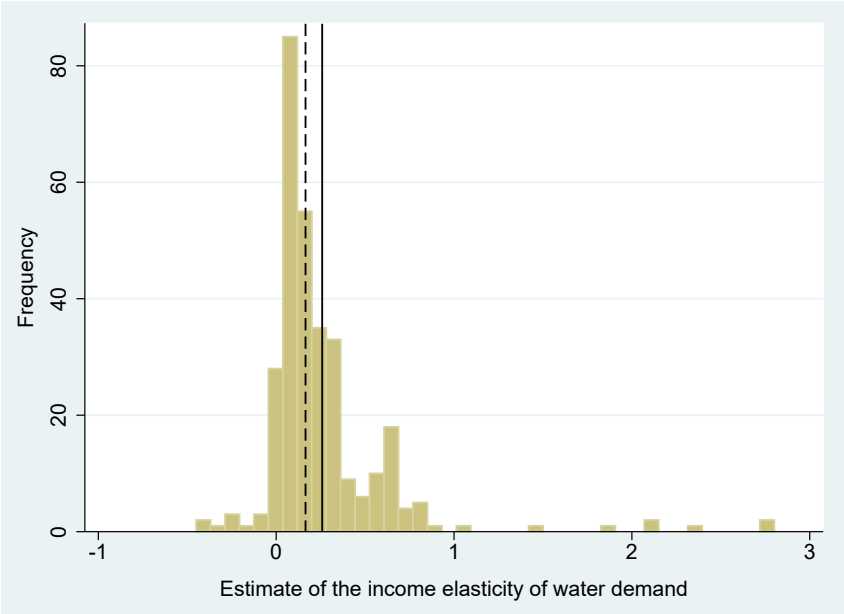
Saleth & Dinar (1997) do not use the straightforward income variable in the demand function but instead proxy for income using housing categories. Schefter & David (1985) estimate the level-level linear functional form of (1), while Gibbs (1978) and Jones & Morris (1984) estimate the semi-log functional form of (1). Gaudin *et al.* (2001) and Nieswiadomy (1992) do not report standard errors for their estimates.

Our final data sample comprises 307 income elasticity estimates taken from 62 studies listed in Table 1. The oldest study was published in 1972 and the most recent one in 2015, which means that this meta-analysis covers as long a period of time as the two previous meta-analyses combined. The apparently right-skewed distribution of estimates is shown in Figure 1. The reported elasticities range from -0.45 to 2.8 and are characterized by a mean of 0.26 and a median of 0.16. Less than 3% of the estimates are larger than 1, which suggests that the demand for water is inelastic with respect to income. More than 94% of the estimates are higher than 0, which supports the intuition that water is not an inferior good. The double-peakedness of the histogram indicates the presence of systematic heterogeneity in the estimates; moreover, Figure 2 reveals the presence of substantial within- and between-study variation. Consequently, for each estimate we collect more than 30 explanatory variables describing the characteristics of the estimation models and investigate the possible reasons for heterogeneity in Section 4.

Table 1: Studies used in the meta-analysis

Agthe & Billings (1980)	Gaudin (2005)	Nieswiadomy & Molina (1991)
Al-Najjar *et al.* (2011)	Gaudin (2006)	Olmstead (2009)
Al-Qunaibet & Johnston (1985)	Hanemann & Nauges (2005)	Olmstead *et al.* (2007)
Asci & Borisova (2014)	Hewitt (1993)	Piper (2003)
Ayadi *et al.* (2002)	Hewitt & Hanemann (1995)	Polycarpou & Zachariadis (2013)
Bartczak *et al.* (2009)	Hoffmann *et al.* (2006)	Reynaud *et al.* (2005)
Basani *et al.* (2008)	Hoglund (1999)	Rietveld *et al.* (1997)
Billings (1982)	Horn (2011)	Schleich & Hillenbrand (2009)
Billings & Agthe (1980)	Hussain *et al.* (2002)	Sebri (2013)
Binet *et al.* (2012)	Jia & Bao (2014)	Statzu & Strazzera (2009)
Binet *et al.* (2014)	Lyman (1992)	Strand & Walker (2005)
Carter & Milon (2005)	Mansur & Olmstead (2012)	Strong & Smith (2010)
Cheesman *et al.* (2008)	Miyawaki *et al.* (2011)	Tabieh *et al.* (2012)
Dalmas & Reynaud (2004)	Monteiro & Roseta-Palma (2011)	Taylor *et al.* (2004)
Darr *et al.* (1975)	Musolesi & Nosvelli (2007)	Williams (1985)
Dharmaratna & Parasnis (2011)	Mylopoulos *et al.* (2004)	Williams & Suh (1986)
Fenrick & Getachew (2012)	Nauges & Strand (2007)	Wong (1972)
Foster & Beattie (1979)	Nauges & Thomas (2003)	Yoo (2007)
Foster & Beattie (1981)	Nauges & Van Den Berg (2009)	Younes & Matoussi (2011)
Frondel & Messner (2008)	Nieswiadomy (1992)	Zapata (2015)
Garcia & Reynaud (2004)	Nieswiadomy & Cobb (1993)	

Figure 1: The histogram suggests substantial heterogeneity and under-reporting of negative estimates

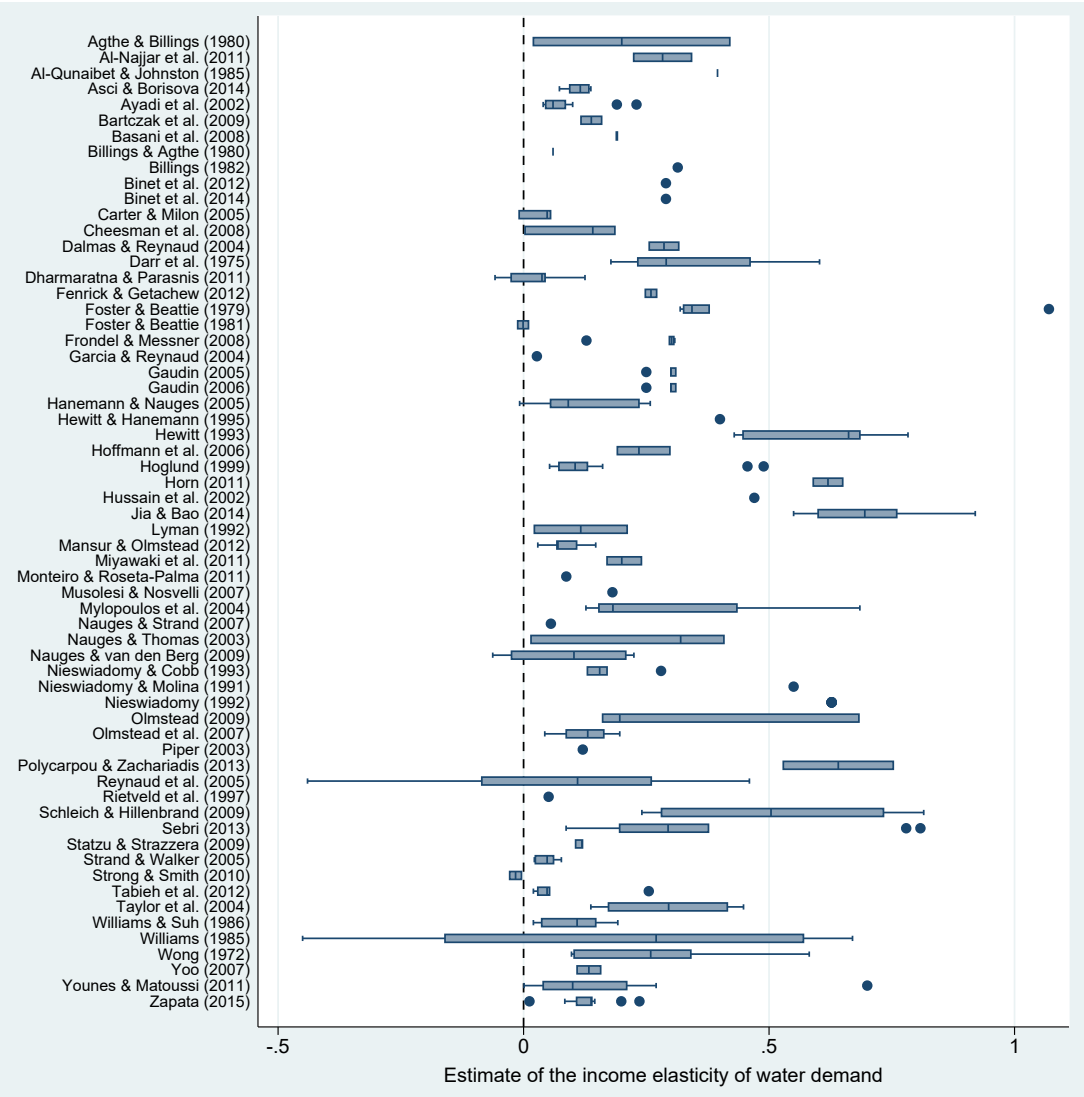


Notes: The figure depicts a histogram of the estimates of the income elasticities of water demand reported by individual studies. The dashed vertical line denotes the sample median; the solid vertical line denotes the sample mean.

To gain a first insight into the potential causes of heterogeneity we compute mean values of the income elasticity estimates for different groups of data, methods, and publication characteristics. Table 2 reports the results for both unweighted estimates and estimates weighted by the inverse of the number of estimates per study, such that studies with many estimates do not drive the mean. We show that short-run elasticities are, on average, 0.1 larger than long-run elasticities. The difference, however, disappears when we give each study the same weight; hence, we do not further divide our sample between short- and the long-run elasticities but analyze the pooled data set while controlling for this difference (which is in accordance with the approach of the previous meta-analyses Dalhuisen *et al.*, 2003; Sebri, 2014). On the one hand, studies using data aggregated at the municipal level yield nearly identical estimates to studies that employ individual household data. On the other hand, the difference between the published and unpublished studies is robust to weighting and fluctuates around 0.1. Differences in results based on publication outlet often, although not necessarily, indicate the presence of publication bias in the literature, as we will discuss below.

Table 2 further suggests that the reported elasticities vary across countries. First, the mean elasticities are consistently higher for the United States, lower for Europe, and even lower for

Figure 2: Estimates of the elasticity vary within and across studies



Notes: The figure shows a box plot of the estimates of the income elasticity of water demand reported in individual studies. Outliers are excluded from the figure but included in all statistical tests.

Table 2: Income elasticity estimates for different subsets of data

		Unweighted			Weighted		
	No. of observations	Mean	95% conf. int.		Mean	95% conf. int.	
Temporal dynamics							
Short-run elasticity	216	0.291	0.233	0.348	0.274	0.221	0.326
Long-run elasticity	91	0.189	0.149	0.229	0.251	0.204	0.299
Aggregation level							
Household data	194	0.254	0.196	0.311	0.290	0.232	0.347
Aggregate data	113	0.273	0.214	0.332	0.235	0.184	0.286
Publication status							
Unpublished studies	66	0.347	0.212	0.482	0.366	0.241	0.492
Published studies	241	0.237	0.198	0.277	0.251	0.210	0.292
Spatial variation							
US	136	0.324	0.239	0.410	0.327	0.243	0.411
Europe	51	0.261	0.202	0.321	0.251	0.195	0.307
Other than US or Europe	120	0.188	0.149	0.227	0.212	0.172	0.251
Developed countries	201	0.295	0.235	0.356	0.291	0.234	0.348
Developing countries	106	0.195	0.154	0.236	0.219	0.177	0.262
Estimation technique							
No endogeneity control	142	0.268	0.213	0.322	0.282	0.225	0.340
Endogeneity control	165	0.255	0.191	0.319	0.260	0.203	0.317
All estimates	307	0.261	0.218	0.303	0.269	0.228	0.309

Notes: The table reports mean values of the income elasticity estimates for different subsets of data. The exact variable definitions are available in Table 4. Weighted = estimates are weighted by the inverse of the number of estimates per study.

the rest of the countries in our sample. Second, the mean elasticities show that the level of development matters: studies of developed countries report estimates that are 0.1 higher than studies of developing countries, on average. This result is, however, counter-intuitive: one would expect that households from developing countries would more vigorously use the opportunity to consume more water when they are able to afford to do so compared to households from developed countries. This result might be explained by different expenditure structures. The income elasticity in developed countries may be higher since water in some cases becomes a luxury good (used for filling up swimming pools, washing cars, and watering lawns). Similarly, the income elasticity in developing countries may be lower since a significantly higher proportion of income goes to other necessities, such as food or clothing.

A common problem associated with demand equations with block rates is that prices are endogenously determined by the quantity demanded. Therefore, researchers using estimation techniques that do not account for endogeneity violate the assumption of no correlation between the explanatory variables and the error term. Some authors acknowledge the prob-

lem and attempt to justify their 'inappropriate' method choice (for example Foster & Beattie, 1979, disregard endogeneity due to nature of their data set), but few test for simultaneity, as in Nieswiadomy & Molina (1989) using the Hausman (1978) test or Williams (1985) using a Ramsey-type test. Some researchers even argue that given the similarity of estimates produced by OLS (not accounting for endogeneity) and the use of instruments (IV-based techniques accounting for endogeneity), simple OLS might suffice for demand analyses under block-rate pricing (see the detailed methodological survey of Arbues *et al.*, 2003, who follow the arguments originally advanced by Saleth & Dinar, 1997).

Our comparison of the reported income elasticities between the estimation techniques accounting for price endogeneity (such as the IV method or the generalized method of moments) and estimation techniques disregarding endogeneity (generally OLS and random effects) appears to be consistent with that of Arbues *et al.* (2003): based on the simple and weighted means from Table 2 we do not observe any large differences between the estimation techniques. Although Arbues *et al.* (2003) mentions that OLS under different block tariffs may underestimate or overestimate demand elasticity depending on whether the supply schedule is steeper than the demand schedule, based on our simple analysis one would argue that estimated elasticities do not depend on whether a researcher addresses the problem of endogeneity. This conclusion would be in line with previous meta-analyses on elasticities of water demand (Espey *et al.*, 1997; Dalhuisen *et al.*, 2003; Sebri, 2014), which do not find significant dependencies between the different estimation techniques and the estimated water demand elasticities.

None of the previous meta-analyses, however, has tested for publication selection. Publication bias, if present, can seriously distort the picture offered by the literature (Doucouliagos & Stanley, 2013). For example, Stanley (2005), correcting the results of Dalhuisen *et al.* (2003) for publication bias, finds the estimates of the price elasticity of water demand to be exaggerated fourfold. Ashenfelter *et al.* (1999), who test for publication bias in estimates of the schooling-earnings relationship, report that publication bias plagues the IV estimates, which typically yield higher standard errors (and thus researchers search for systematically higher estimates to achieve the desired level of statistical significance). It follows that the comparison of sample averages sheds some light on the sources of the heterogeneity of the estimates, but it does not reflect the differences in the underlying elasticity if the estimates are subject to publication bias.

3 Detecting Publication Bias

Publication bias arises when some estimates have a higher probability of being reported than other estimates. Researchers may prefer to report strong (i.e., statistically significant) and useful findings that tell a good story; editors and referees may prefer significant findings that are in line with theory. Theoretically, since water does not have any close substitute, once the household's income increases, the demand for water should also increase. Therefore, water cannot be considered an inferior good, and its income elasticity of demand should be positive. Given the strong case for positive estimates, it is not surprising that researchers treat negative estimates with suspicion. Hence Rietveld *et al.* (1997, p. 30) comment on their estimated income elasticities as follows: "*The results are terrible... parameters are having the 'wrong' [negative] sign...*" The previous meta-analyses on elasticities of water demand (Espey *et al.*, 1997; Dalhuisen *et al.*, 2003; Sebri, 2014) call the price elasticities with the unintuitive sign *'perverse'* and eliminate them from their samples.

But even by the law of chance, negative estimates of income elasticity should occasionally appear in the literature. The probability of negative estimates increases with small samples, noisy data, or misspecification of the demand function (more in Stanley, 2005). Consequently, researchers tend to suppress their negative estimates; such a practice would, even if beneficial at the level of individual studies, drive the global mean of the reported elasticities upwards. Doucouliagos & Stanley (2013) find that most fields of economic research are affected by publication selection bias. The field of energy and resource economics research is no exception: Havranek *et al.* (2012) and Havranek & Kokes (2015) find publication bias in the literature estimating the price and income elasticities of gasoline demand, while Havranek *et al.* (2015b) report the same problem in the literature on the social cost of carbon.[1]

The most common visual tool used for the investigation of the presence of publication bias is the funnel plot (Egger *et al.*, 1997). Figure 3 depicts the plot for all 307 estimates of the income elasticity of water demand on the horizontal axis and the inverse of the standard error of an estimate used as a measure of precision on the vertical axis. Ideally, the plot should resemble an inverted funnel: the estimates with the highest precision should be close to the true effect, while

[1] Examples of publication bias in other fields include Havranek & Irsova (2011) and Havranek & Irsova (2012) on foreign direct investment spillovers, Havranek (2015) on the elasticity of intertemporal substitution, and Havranek *et al.* (2016) on the natural resource curse.

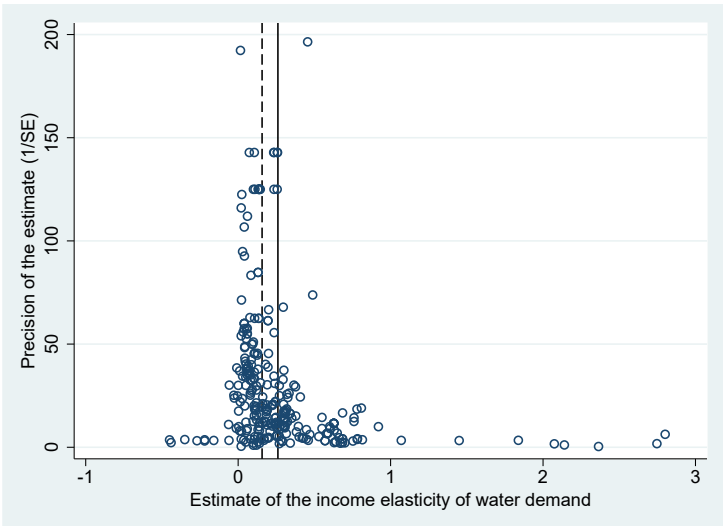
the estimates with decreasing precision are more dispersed from the mean (Havranek & Irsova, 2017). If publication bias is present, the funnel is asymmetrical (when the bias is related to the sign of the effect) and hollow and wide (when the bias is related to the significance of the effect). In Figure 3 we observe that the left-hand part of the funnel is essentially absent. Researchers indeed omit negative values of the elasticity, which biases the mean reported estimate upwards.

We support our conclusions from the funnel plot using a more formal analysis following Stanley (2005), who examines the correlation between the estimates and their standard errors:

$$YED_{ij} = YED_0 + \beta \cdot SE(YED_{ij}) + \mu_{ij}, \tag{2}$$

where YED_{ij} denotes i-th effect and its standard error $SE(YED_{ij})$ estimated in the j-th study, and μ_{ij} is the error term. The intercept of the equation, YED_0, is the true mean elasticity beyond publication bias. If no publication bias is present in the literature, the coefficient β should be zero (the methods used by researchers imply that the ratio of the point estimate to the standard error has a t-distribution, which means that the two variables should form statistically independent quantities). Otherwise, we should observe that the estimated effects are correlated with their standard error, for example because researchers with large standard errors need large point estimates to produce statistical significance, or because they discard negative estimates, which yields a positive β due to the heteroskedasticity of (2).

Figure 3: The funnel plot suggests publication bias



Notes: The dashed vertical line indicates the median estimate of the income elasticity of water demand; the solid vertical line indicates the mean estimate of the income elasticity of water demand. When there is no publication selection bias, the estimates should be symmetrically distributed around the mean effect.

Equation (2) can be presented as a funnel asymmetry test, as it follows from rotating the axes of the funnel plot and inverting the value of precision to display the standard error. To account for heteroskedasticity and within-study dependence in (2), we report robust standard errors clustered at the study level. Further, we estimate different specifications: 1) the original unweighted data sample, 2) weighting by the inverse of the number of estimates per study (small and large studies are thus given the same importance), and 3) weighting by the inverse of the standard error (precise estimates are given greater weight). We estimate each specification using simple OLS with study-level fixed effects to account for unobserved study-level characteristics.

Table 3 presents the results of the funnel asymmetry tests. In Panel A of Table 3 we show the different specifications of (2) applied to the full sample of 307 elasticity estimates. The results corroborate the findings from the funnel plot that publication selection bias is present in the literature on the income elasticity of water demand. The results from Panel A also place the true effect in the literature at approximately 0.178, which means that increasing a household's income by one percent increases water demand by 0.18 percent. This value is fairly robust throughout different estimations in Panel A (with one exception in the last column; nevertheless, the combination of precision weighing and study fixed effects often produces unstable results). The coefficient corresponding to publication bias has a positive sign, which means that the true effect is probably smaller than what researchers tend to report on average. Our estimate of the effect is relatively close to the mean estimate from Sebri (2014), who argue the number to be half the mean reported by Dalhuisen *et al.* (2003).

As a complementary analysis, we show another visual test, the Galbraith plot, which focuses on the bias caused by the preference for significant results. Authors who prefer significant results and disregard insignificant estimates will over-report high t-values (in absolute terms). We follow Stanley (2005), Havranek (2010), and Havranek *et al.* (2018) and define the standardized t-statistics $T(YED_{ij})$ adjusted for the true effect from Table 3:

$$T(YED_{ij}) = \frac{YED_{ij} - YED_0}{SE(YED_{ij})}, \tag{3}$$

where YED_0 represents the true effect estimated by the funnel asymmetry test, and YED_{ij} represents the i-th estimate of the income elasticity with $SE(YED_{ij})$ as the corresponding standard error reported in the j-th study. For YED_0, we employ the baseline true effect from the first column of Panel A in Table 3, 0.178, and plot the final statistics in Figure 4.

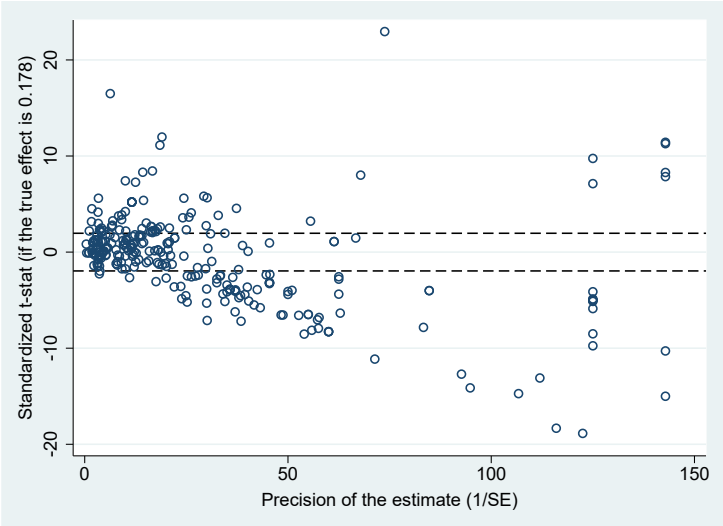
Table 3: Tests show publication bias in estimates that control for endogeneity

Panel A: Whole Sample	Unweighted		Study		Precision	
	OLS	FE	OLS	FE	OLS	FE
SE (publication bias)	0.676**	0.551*	0.884***	0.644***	1.280***	1.514
	(0.305)	(0.301)	(0.132)	(0.161)	(0.369)	(1.176)
Constant (effect beyond bias)	0.178***	0.193***	0.155***	0.187***	0.103***	0.121***
	(0.029)	(0.037)	(0.022)	(0.021)	(0.012)	(0.045)
Observations	307	307	307	307	307	307
Panel B: No Endogeneity Control	Unweighted		Study		Precision	
	OLS	FE	OLS	FE	OLS	FE
SE (publication bias)	0.290	0.286	0.753	0.523	1.010**	0.326
	(0.307)	(0.288)	(0.576)	(0.450)	(0.405)	(0.340)
Constant (effect beyond bias)	0.223***	0.224***	0.188***	0.218***	0.112***	0.148***
	(0.0347)	(0.0445)	(0.0569)	(0.0581)	(0.0121)	(0.0153)
Observations	142	142	142	142	142	142
Panel C: Endogeneity Control	Unweighted		Study		Precision	
	OLS	FE	OLS	FE	OLS	FE
SE (publication bias)	1.053***	1.054**	0.919***	0.834***	1.650***	3.689***
	(0.252)	(0.437)	(0.0942)	(0.166)	(0.490)	(1.159)
Constant (effect beyond bias)	0.153***	0.153***	0.140***	0.151***	0.0959***	0.0631
	(0.0327)	(0.0421)	(0.0246)	(0.0217)	(0.0153)	(0.0396)
Observations	165	165	165	165	165	165

Notes: The table reports the results of the regression $YED_{ij} = YED_0 + \beta \cdot SE(YED_{ij}) + \mu_{ij}$, where YED_{ij} denotes i-th effect estimated in j-th study, and $SE(YED_{ij})$ denotes its standard error, estimated either by ordinary least squares (OLS) or study-level fixed effects (FE). Panel A reports results for the full sample of estimates, Panel B reports the results for the subset of elasticities computed by estimation methods not accounting for price endogeneity in the demand function, and Panel C reports the results for the sample where the estimation methods do account for endogeneity in the demand function. Unweighted = model is not weighted; Study = model is weighted by the inverse of the number of estimates per study; Precision = model is weighted by the inverse of the standard error of an estimate. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors, in parentheses, are clustered at the study level.

Figure 4 represents a Galbraith plot, a scatter plot with the precision of an estimate $1/SE(YED_{ij})$ on the horizontal axis against the standardized size of the t-statistics $T(YED_{ij})$ on the vertical axis (Galbraith, 1990). Although a large number of estimates are situated between the two lines denoting the critical values of the t-statistic for the 5% significance level, the plot indicates some publication bias since the number of the estimates outside the area defined by the two dashed lines increases with precision. Figure 4 also indicates excess variation in the standardized t-values since only 43% of the estimates fall into the area where they should be given the properties of the t-statistic and the correctness of our estimate of the underlying elasticity. It follows that researchers are more likely to prefer significant results over insignificant results and possibly conduct a specification search to produce the desired outcomes.

Figure 4: Galbraith plot suggests publication bias



Notes: The horizontal dashed lines form the boundary of the $(-1.96, 1.96)$ interval, which should not be surpassed in more that 95% of cases if there is no publication bias related to statistical significance and no heterogeneity.

We have established thus far that, overall, the literature on the income elasticity of water demand suffers from publication bias related to the sign and significance of the estimates. The contamination of the literature by publication bias compels us to reassess the conclusions drawn from comparing the average reported estimates for different subsamples with respect to data and method choices. Especially, as Ashenfelter *et al.* (1999) note, researchers are often more likely to report larger estimates to compensate for the large standard errors when instrumental variables are employed. We show in Table 2 that, on average, the methods controlling for endogeneity and not controlling for endogeneity do not yield any notable differences in elasticities. Following Ashenfelter *et al.* (1999), we investigate whether there is selective reporting related to the method choice that might drive the publication bias in the literature.

For the analysis, let us return to Table 3. Panels B and C show the funnel asymmetry tests applied to two groups of estimates: those that do not control for endogeneity (Panel B) and those that do control for endogeneity (Panel C). We demonstrated in the previous section that the mean reported estimates are very similar for both groups. The similarity disappears, however, when we account for publication bias. We find no bias for estimates that do not control for endogeneity, and the underlying elasticity for these estimates is approximately 0.22. Regarding the endogeneity-consistent estimates, however, we find evidence of *substantial* publication bias

in the classification due to Doucouliagos & Stanley (2013). Here, the mean estimate is therefore biased upwards, and the underlying elasticity is only 0.15 or less. Thus, both endogeneity and publication biases matter. The studies that control for endogeneity bias would tend to report estimates that are substantially smaller than OLS studies if they were not more susceptible to publication selection. Publication bias in the better-specified studies arises because researchers need larger estimates to offset large standard errors—or, simply, because researchers care more about IV estimates in the first place and only report OLS results as additional results. There might exist, however, other data and method choices that are also correlated with publication bias or the underlying effect. We address these issues in the next section.

4 Why Do the Estimates Vary?

4.1 Variables and Estimation

Table 2 and Figure 2 in Section 2 present a first tentative examination of the potential sources of heterogeneity behind the estimates of the income elasticity of water demand. To investigate this heterogeneity more systematically, we augment regression (2) by including a plethora of explanatory variables: 30 study design characteristics and the interaction term between the standard error and a dummy variable that equals one if the study in question does not control for price endogeneity. The explanatory variables capturing the variation in data and methodology are listed in Table 4; the table provides the definitions of the variables and their summary statistics, including the simple mean, standard deviation, and the mean weighted by the inverse of the number of observations extracted from a study.

For ease of exposition we divide the estimate and study characteristics into variables reflecting the specification of the demand function (8 aspects), price specification (2 aspects), data characteristics (7 aspects), estimation technique (3 aspects), tariff structure (3 aspects), countries examined (3 countries), and publication characteristics (4 aspects). We note that this section merely serves as a means of discussing the main sources of heterogeneity and not as an exhaustive survey of the methods used in the literature estimating water demand elasticities. For a more detailed discussion we refer the reader to the previous and competently executed meta-analyses of Dalhuisen *et al.* (2003) and Sebri (2014).

Table 4: Description and summary statistics of regression variables

Variable	Description	Mean	SD	WM
Income elasticity	The estimate of the income elasticity of water demand.	0.261	0.377	0.270
Standard error	The standard error of the estimate of the income elasticity of water demand.	0.123	0.232	0.130
SE · No endog. control	Interaction term between the standard error and the estimation methods not addressing price endogeneity.	0.071	0.178	0.053
Water demand specification				
Household size	= 1 if the demand equation controls for household size (usually defined as a number of persons living in a household).	0.518	0.500	0.533
Population density	= 1 if the demand equation controls for population density (which often serves as a proxy for lawn size).	0.107	0.310	0.099
Temperature	= 1 if the demand equation controls for temperature.	0.489	0.501	0.427
Rainfall	= 1 if the demand equation controls for rainfall.	0.632	0.483	0.535
Evaporation	= 1 if the demand equation controls for evaporation.	0.130	0.337	0.161
Difference variable	= 1 if the demand equation contains the variable accounting for the difference between the water bill priced at actual rates and the water bill priced at marginal prices (Dalhuisen *et al.*, 2003).	0.156	0.364	0.218
Lagged dep. variable	= 1 if the demand equation contains the lagged dependent variable.	0.085	0.279	0.124
Discrete-continuous	= 1 if the demand equation is based on the discrete-continuous model.	0.107	0.310	0.125
Price specification				
Marginal price	= 1 if marginal price computed as the price of the last cubic meter of water is used for estimation (reference category for this group of dummy variables: average price computed as the total bill divided by total consumption).	0.401	0.491	0.462
Other price	= 1 if a price other than marginal or average is used for estimation (such as the Shin price deployed by Shin, 1985).	0.130	0.337	0.198
Data characteristics				
Long-run elasticity	= 1 if the estimated elasticity is the long-term instead of short-term elasticity.	0.296	0.457	0.237
Household data	= 1 if residential data are used for the estimation instead of data aggregated at the municipal level (including residential, industrial, and commercial water demand).	0.632	0.483	0.597
Daily data	= 1 if the frequency of data used for estimation is daily instead of quarterly, monthly, or annual.	0.189	0.392	0.161
Monthly data	= 1 if the frequency of data used for estimation is monthly instead of daily, monthly, or annual.	0.394	0.489	0.483

Continued on next page

Table 4: Description and summary statistics of regression variables (continued)

Variable	Description	Mean	SD	WM
Annual data	= 1 if the frequency of data used for estimation is annual instead of daily, monthly, or quarterly.	0.235	0.424	0.242
Cross-section	= 1 if cross-sectional data are used for estimation instead of time-series or panel data.	0.293	0.456	0.334
Time-series	= 1 if time series data are used for estimation instead of cross-section or panel data.	0.029	0.029	0.086
Estimation technique				
No endogeneity control	= 1 if the estimation method does not account for price endogeneity; typically ordinary least squares (reference category for this group of dummy variables is the use of instrumental variables).	0.463	0.499	0.411
Panel technique	= 1 if a fixed effects panel technique is employed for estimation.	0.244	0.430	0.212
Other estimator	= 1 if an estimation method accounting for endogeneity other than instrumental variables and panel fixed effects is employed for estimation.	0.111	0.314	0.208
Tariff structure				
Flat tariff	= 1 if a flat tariff structure is used for estimation (reference category for this group of dummy variables is the situation in which the tariff structure employed is not available).	0.078	0.269	0.121
Increasing tariff	= 1 if an increasing tariff structure is used for estimation.	0.485	0.501	0.526
Decreasing tariff	= 1 if a decreasing tariff structure is used for estimation.	0.023	0.150	0.031
Countries examined				
Europe	= 1 if the income elasticity of water demand is estimated for a location in Europe, instead of the US or other countries.	0.166	0.373	0.226
Other location	= 1 if the income elasticity of water demand is estimated for other location than Europe or the US.	0.391	0.489	0.355
Developed countries	= 1 if the income elasticity of water demand is estimated for a developed country instead of developing country.	0.655	0.476	0.693
Publication characteristics				
Publication year	The publication year of the study (the base year is the sample minimum: 1972).	30.29	11.36	30.02
Citations	The average yearly number of citations the study received in Google Scholar since its appearance there.	4.494	6.724	4.385
Impact factor	RePEc recursive discounted impact factor for journals.	0.106	0.199	0.088
Published	= 1 if the study is published in a peer-reviewed journal.	0.799	0.416	0.839

Notes: SD = standard deviation, SE = standard error, WM = mean weighted by the inverse of the number of estimates reported per study.

Water demand specification. Researchers specify the water demand equation to reflect the behavioral patterns of consumers under certain living conditions. We codify several of these patterns and conditions as the possible sources of heterogeneity. For example, we consider a dummy for including a control for *household size* (the number of people living in a household) since, due to economies of scale, individual consumption should decrease with an increase in household size (Arbues *et al.*, 2010). We also take into account whether the authors include *population density* in their demand equation, which is often used as a proxy for the housing stock and size of yards (Gaudin, 2005). Some authors include the *difference variable*, which reflects the difference in the actual water bill and the water bill priced at marginal prices (Espey *et al.*, 1997), or as Dalhuisen *et al.* (2003) call it, a lump-sum transfer imposed by the tariff structure. Moreover, Hewitt & Hanemann (1995) suggest using the *discrete-continuous model* to account for the discrete price structure of water tariffs and the continuous consumption of water. Dalhuisen *et al.* (2003) report the inclusion of a *difference variable* to increase and the choice of a *discrete-continuous model* to decrease the income elasticities.

Dalhuisen *et al.* (2003) show that water demand is sensitive to weather factors. Some authors (like Miaou, 1990) criticize the assumption of a linear relationship between water demand and weather variables, suggesting that rainfall might have a dynamic effect on water consumption, and investigate the possibility of a threshold beyond which precipitation or temperature does not affect water use. We use dummy variables for studies that include information on *temperature*, *rainfall*, and *evaporation*. Dalhuisen *et al.* (2003) reveal that authors using weather-conditioning variables report higher income elasticities when evaporation is included in the model while Sebri (2014) shows the inclusion of rainfall variable decreases the elasticities. We also code for whether the study makes a dynamic adjustment of the demand model with a *lagged dependent variable*, which mostly reflects the fact that water use is a habit and that time is required to change this habit in response to other, usually price or weather, changes (Asci & Borisova, 2014). It is worth noting, however, that the inclusion of the lagged dependent variable in the demand model can violate the assumptions of some simple estimation techniques.

Price specification. Water is also considered to be inelastic in price because consumers typically exhibit limited awareness of the pricing structure. The suitability of using the *average*, the *marginal*, or *other* pricing schemes in the water demand function remains a matter of heated

discussion. On the one hand, Nauges & Van Den Berg (2009), among others, argue that since consumers are rarely aware of their rate structure, they react to their average bill rather than to their marginal bill (which conforms to our expectations). On the other hand, Saleth & Dinar (1997) argue that the use of marginal pricing including the difference variable instead presupposes average price behavior and has many methodological advantages. Shin (1985) introduces a price-perception concept that identifies which of the two prices (the marginal or the average price) is better understood by consumers. Few researchers have followed in his footsteps; the recent work by Binet *et al.* (2014) proposes significant modifications to the functional form of Shin's perceived price. The following recent papers also provide additional insight into this issue: Ito (2013), Wichman (2014), and Wichman *et al.* (2016). The reference category for this group of dummy variables is the average pricing scheme.

Data characteristics. Given the small differences between the averages of the short-run and *long-run elasticities* found in Table 2, we do not divide the sample accordingly, but we still control for this form of temporal dynamics in our model. If significant, we would expect higher responsiveness to changes in income in the longer time period. Moreover, we take into account whether the study uses *household data* only or aggregates the data at the municipal level, including household, industrial, and commercial water consumption since the households could prove to be more sensitive to changes in income than the rest of the water consumers. Although the residential elasticity is considerable more important to us, given the relatively small number of observations in this study, we also account for the aggregated estimates and err on the side of inclusion.

Another characteristic we focus on is the frequency of the data: higher frequencies provide less-detailed information on immediate behavioral patterns, and although water is inelastic in income, the granularity (in our case *yearly*, *quarterly*, *monthly*, and *daily*) also matters in the previous meta-analyses. Sebri (2014), for example, documents that higher data frequencies deflate the elasticities. The reference category for the data frequency is the use of quarterly data for estimation. We also distinguish among *time series*, *cross-sectional* data, and panel data, using panel data as the reference category. The previous meta-analyses report that the time dimension of the data systematically adds to an increase in the elasticity estimates.

Estimation technique. The most commonly used estimation techniques are ordinary least squares (Nieswiadomy & Molina, 1991), two-stage least squares (Nieswiadomy & Molina, 1991), three-stage least squares (Al-Najjar *et al.*, 2011), generalized method of moments (Musolesi & Nosvelli, 2007), and panel techniques with random or fixed effects (Cheesman *et al.*, 2008; Sebri, 2013). We mark all techniques that do not account for the price endogeneity present in the demand equation as *No endogeneity control.*[2] Given that we find no publication bias in the estimates produced by methods that ignore endogeneity, we hypothesize the endogeneity variable and the interaction term between the standard error and the endogeneity variable to be significant. The reference category for this group of dummy variables is the instrumental variables estimation method and its derivatives.

Tariff structure. Tariff structures help policy makers control the demand for water. An increasing structure, for example, means that the price is constant within discrete intervals of use but increasing between the different intervals of use. The outcomes of such water policies are, however, not always clear-cut. In the case of the *increasing tariff* structure, the policy is expected to limit excessive consumption of water. This leads to higher real income, and if the income elasticity of water demand is positive, higher real income results in higher demand for water. It is unclear, however, which of these two effects prevails. To address such problems, we include information on the use of *flat*, *increasing*, and *decreasing tariff* structures. The reference category for this group of dummy variables is the situation in which the tariff structure employed is not available.

Countries examined. The main reasons for cross-country heterogeneity are potential differences in consumption habits, culture, climate, and path-dependency in policy. The previous meta-analyses are rather inconclusive with respect to spatial variation: while Dalhuisen *et al.* (2003) find a significant difference between income elasticity estimates for the US and Europe, Sebri (2014) argues that this difference is insignificant. Hence, we distinguish among different locations in the *US*, *Europe* (including Cyprus, France, Germany, Greece, Italy, Poland,

[2]Note that if a flat tariff rate is imposed, the (constant) price of water is exogenous to water demand, and thus, the endogeneity problem does not need to be addressed. As there are only 11 such observations of the elasticity for a flat tariff structure estimated by OLS, we treat them as any other observation of the elasticity estimated by techniques not accounting for endogeneity. Robustness checks, in which these estimates are eliminated, yield very similar conclusions to those in Table 3 and Table 5.

Portugal, Slovakia, and Sweden), and any location outside the US and Europe (such as Australia, Cambodia, Canada, China, Ecuador, Indonesia, Israel, Japan, Jordan, Korea, Kuwait, Sri Lanka, Tunisia, and Vietnam). The reference category for this group of dummy variables is the estimation of the income elasticity of water demand for a location in the US.

Furthermore, we distinguish between whether the study estimates the elasticity for a *developed country* or a *developing country*. The inhabitants of developing countries are forced to consume a lower amount of water since they typically not have sufficient income to be able to afford more; changes in income may thus have different effects in these countries compared to developed countries. Similarly, we assume the water consumption of individuals living in developed countries to be sufficient; hence, a change in income should not trigger a significant change in water consumption. Altogether, individuals from developed countries are expected to dedicate a relatively lower proportion of their additional income to expenditures on water than individuals from developing countries.

Publication characteristics. We employ several publication characteristics as proxies for methodological advances that might not be directly captured by our methodological variables. For example, the variable *publication year* could tell us whether newer studies tend to report systematically different elasticities. To address the quality of a study, we use the average yearly *number of citations* and the RePEc recursive discounted *impact factor* for journal publications. We also distinguish between *published* (journal publications) and *unpublished* studies (working papers and other unrefereed materials) since the previous meta-analyses also lack consensus on this matter: while Dalhuisen *et al.* (2003) find that published estimates of elasticities are smaller than the unpublished ones, Sebri (2014) finds the opposite.

Our intention is to examine whether the evidence for publication bias remains strong if we control for the possible causes of heterogeneity. Ideally, we would like to regress the collected income elasticities of water demand on all of the explanatory variables at hand (like Dalhuisen *et al.*, 2003; Sebri, 2014, do). Given that we have so many variables, however, some of them will likely be insignificant, which would inflate the variation of other estimated parameters in the regression and introduce inefficiency. Alternatively, sequential t-tests can be employed, but eliminating insignificant variables one by one might lead to a loss of important variables during the process. Following Havranek & Irsova (2017) we instead employ Bayesian model

averaging (BMA), which formally addresses such model uncertainty. BMA goes through millions of different models created from the subsamples of the potential explanatory variables and searches for those models with the highest explanatory power. In a Bayesian setting, the model's explanatory power is represented by the posterior model probability, which is analogous to the adjusted coefficient of determination in frequentist econometrics. Applications of BMA in meta-analysis include Irsova & Havranek (2013), Havranek *et al.* (2015a), and Havranek *et al.* (2017).

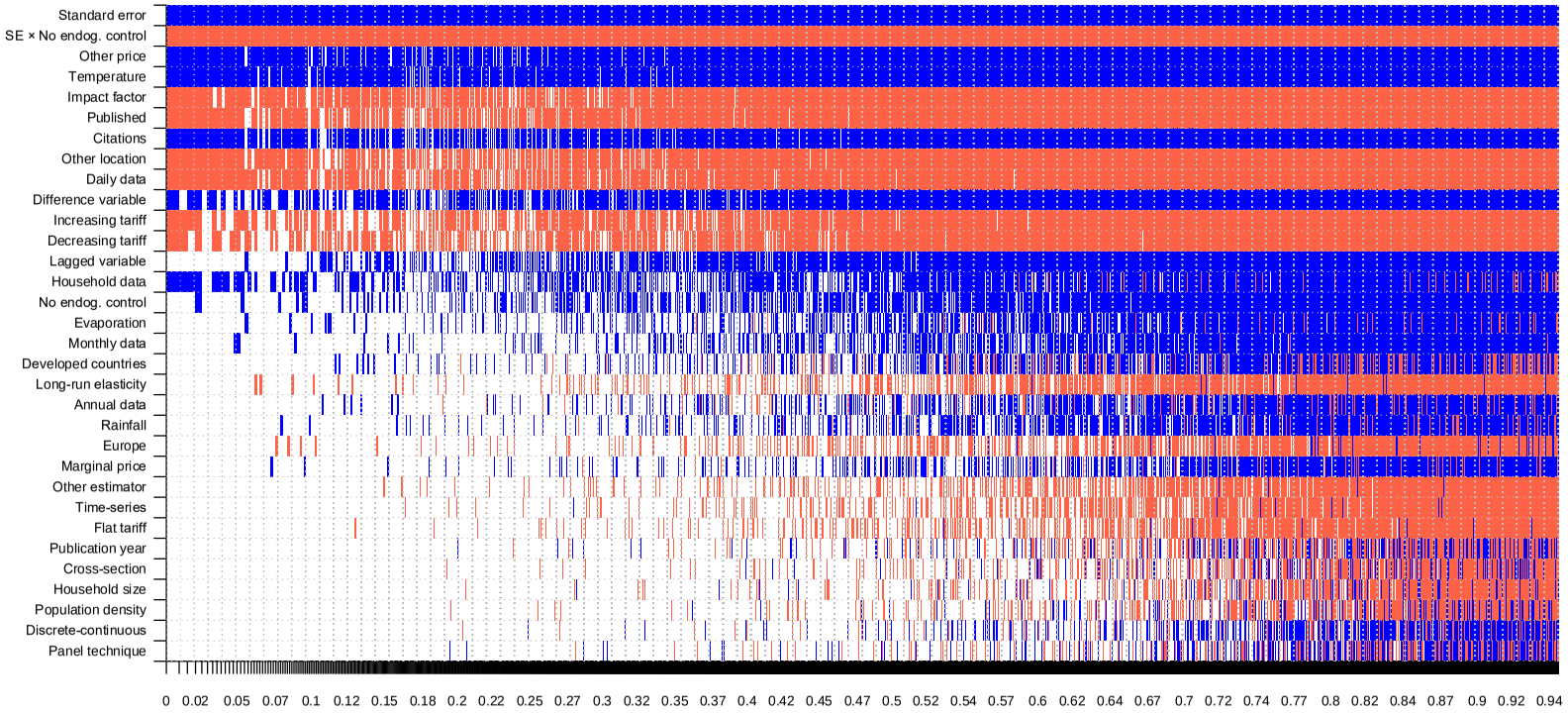
Since we use the `bms` package in `R` (Feldkircher & Zeugner, 2012), our BMA does not estimate all of the 2^{32} possible combinations of models but uses Markov Chain Monte Carlo samplers that propose the candidate models to be estimated (estimating all of the models would take several months). The estimated BMA coefficients, posterior means, are averages of the coefficients across all of the models, weighted by the posterior model probability. Thus, each coefficient has an approximately symmetrical distribution with a posterior standard deviation, which is analogous to the standard error in frequentist econometrics. Each coefficient is assigned a posterior inclusion probability, the sum of posterior model probabilities from all of the models in which the variable is found, which is analogous to statistical significance in frequentist econometrics. Further details on BMA can be found, for example, in Eicher *et al.* (2011).

4.2 Results

Figure 5 presents the results of the BMA exercise. The columns represent different models and are sorted by posterior model probability in descending order from left to right. The rows represent different variables and are sorted by posterior inclusion probability in descending order from top to bottom. Each cell thus belongs to a particular variable in a particular model: if the cell is blue (darker in grayscale), the coefficient of a variable is positive; if the cell is red (lighter in grayscale), the coefficient is negative; if there is no color, the variable is excluded from the model. We observe that almost half of the variables are included in the best model, and the sign of these variables is robust across different models.

The numerical results of the BMA exercise are reported in Table 5 (we follow Eicher *et al.*, 2011, definitions of parameter and model priors). In addition, we report an OLS regression, which includes 14 explanatory variables recognized by the `bms` package in `R` to form the *top model*. The OLS results are consistent with BMA: the estimated coefficients have the same sign

Figure 5: Model inclusion in Bayesian model averaging



Notes: The figure depicts the results of BMA. On the vertical axis, the explanatory variables are ranked according to their posterior inclusion probabilities from the highest at the top to the lowest at the bottom. The horizontal axis shows the values of cumulative posterior model probability. Blue color (darker in greyscale) = the estimated parameter of a corresponding explanatory variable is positive. Red color (lighter in greyscale) = the estimated parameter of a corresponding explanatory variable is negative. No color = the corresponding explanatory variable is not included in the model. Numerical results are reported in Table 5. All variables are described in Table 4. The results are based on the unweighted specification. The robustness check in which the specification is weighted by the number of estimates per study is consistent with the results of the unweighted specification provided in Table 5. Following the detailed reasoning of Zigraiova & Havranek (2016, p. 28-30), we prefer not to weight our model by the inverse of the standard error because of the many problems with this approach when study-invariant variables are included.

and are similar in magnitude; the significance of the estimated parameters mostly corresponds to the values of the posterior inclusion probability. When interpreting the posterior inclusion probability, we follow Jeffreys (1961), who find evidence of an effect that is *weak* for a value between 0.5 and 0.75, *positive* for a value between 0.75 and 0.95, *strong* for a value between 0.95 and 0.99, and *decisive* for a value higher than 0.99. Therefore, we find weak evidence for the presence of an effect of the variables *Difference variable*, *Daily data*, *Other location*, *Citations*, *Impact factor*, and *Published*; we find positive evidence for the variables *Temperature* and *Other price*; and we find decisive evidence for the *Standard error* and the interaction term *SE* · *No endog. control*.

Publication bias and endogeneity. For only two variables is there decisive evidence that they influence the estimated elasticities: *Standard error* and its interaction with techniques not controlling for price endogeneity *SE* · *No endog. control*. The significance of standard error corresponds to the conclusion that publication bias is present in the literature and the estimated coefficient for publication bias survives the inclusion of data and method heterogeneity. We also confirm that the estimates produced by techniques not accounting for endogeneity suffer less from publication bias. Due to the low posterior inclusion probability, BMA did not recognize the variable *No endog. control* as relevant; however, BMA includes the variable *No endog. control* in the top model, which we use for our frequentist check, where the parameter corresponding to that variable is found to be significant and positive (and is thus in line with the intuition and our analysis from the previous section). When endogeneity is accounted for (variable *No endog. control* = 0), the effect of publication bias after controlling for various sources of heterogeneity is 0.956, only marginally smaller than what is presented in Panel C of Table 3. We conclude that the two variables reflecting publication bias are crucial for explaining the differences between the reported estimates of the income elasticity.

Water demand specification. According to our results, the inclusion of one weather variable can particularly drive the estimated elasticities: authors taking into account the outside *Temperature* find the demand for water to be more income elastic (contrary to those not including the variable, who find the income elasticity to be 0.12 smaller if other factors are held constant). This conclusion contradicts the main results of Dalhuisen *et al.* (2003), who instead

find the inclusion of the evaporation variable to be important, or Sebri (2014), who find that controlling for rainfall drives the results (although the robustness check of Sebri, 2014, is in accordance with Table 5). Higher temperature triggers an increase in the demand for water, which can be addressed by spending a higher proportion of income on water; thus, it is sensible to include the temperature in a demand function.

BMA acknowledges only weak evidence for the importance of the *Difference variable*: the inclusion of this variable increases the differences between the marginal price specification and the average price specification (the frequentist check confirms the significance of its impact). The evidence for the importance of the dynamic model (the inclusion of the *Lagged dependent variable*) is even weaker, and given the results of the robustness check we are inclined to disregard it as a driver of the income elasticity. We do not confirm the previous findings of Sebri (2014), who supports the results of Dalhuisen *et al.* (2003) showing that that the use of the discrete-continuous model has a negative effect on the income elasticity.

Price specification and tariff structure. If a price specification other than average or marginal approach is used in the water demand equation, the income elasticity estimates are on average 0.16 higher, *ceteris paribus*. This result contradicts Dalhuisen *et al.* (2003) and Sebri (2014), who find the differences among the price specifications to be statistically indistinguishable. While BMA determined a very weak impact of different tariff structures on the estimated income elasticity, the frequentist check indicates some systematic dependencies: non-flat tariffs make the demand for water more inelastic and indeed seem to be significantly different from flat and other tariff structures, which is in accordance with theory (a more detailed discussion of the theoretical relationship between tariff structures and income elasticity can be found in Dalhuisen *et al.*, 2001). The choice of a certain tariff structure would then be a suitable policy tool for affecting the elasticity of consumers.

Countries examined. The income elasticity of water demand estimated for a location other than Europe and the US tends to be approximately 0.1 lower when compared to the elasticity estimated for the US. Given that this is the only spatial variation detected in our model, we challenge not only Dalhuisen *et al.* (2003), who find differences between income elasticity estimates for Europe and the US, bus also Sebri (2014), who observes no spatial variation

Table 5: Explaining heterogeneity in the estimates of the income elasticity of water demand

Response variable:	Bayesian model averaging			Frequentist check (OLS)		
Income elasticity	Post. mean	Post. SD	PIP	Coef.	Std. error	p-value
Constant	0.206	NA	1.000	0.245	0.049	0.000
Standard error	0.956	0.130	1.000	1.039	0.259	0.000
SE · No endog. control	-0.572	0.175	0.992	-0.714	0.431	0.097
Water demand specification						
Household size	-0.001	0.008	0.027			
Population density	-0.001	0.013	0.027			
Temperature	0.118	0.074	0.807	0.151	0.055	0.006
Rainfall	0.006	0.027	0.082			
Evaporation	0.018	0.062	0.116			
Difference variable	0.083	0.088	0.543	0.144	0.057	0.011
Lagged dependent variable	0.077	0.108	0.396	0.122	0.101	0.228
Discrete-continuous	0.001	0.014	0.026			
Price specification						
Marginal price	0.004	0.021	0.062			
Other price	0.154	0.092	0.818	0.159	0.074	0.032
Data characteristics						
Long-run elasticity	-0.007	0.028	0.091			
Household data	0.035	0.066	0.271			
Daily data	-0.126	0.130	0.573	-0.167	0.061	0.007
Monthly data	0.009	0.032	0.108			
Annual data	0.008	0.036	0.089			
Cross-section	-0.001	0.009	0.027			
Time-series	-0.007	0.043	0.054			
Estimation technique						
No endog. control	0.021	0.044	0.223	0.084	0.039	0.033
Panel technique	0.000	0.010	0.025			
Other estimator	-0.005	0.025	0.055			
Tariff structure						
Flat tariff	-0.004	0.027	0.048			
Increasing tariff	-0.059	0.072	0.467	-0.099	0.051	0.053
Decreasing tariff	-0.141	0.182	0.445	-0.279	0.127	0.028
Countries examined						
Europe	-0.006	0.027	0.066			
Other location	-0.090	0.090	0.586	-0.084	0.046	0.069
Developed countries	0.006	0.042	0.101			
Publication characteristics						
Publication year	0.000	0.001	0.037			
Citations	0.007	0.007	0.615	0.012	0.004	0.001
Impact factor	-0.232	0.181	0.707	-0.274	0.148	0.065
Published	-0.102	0.089	0.650	-0.172	0.054	0.001
Studies	62			62		
Observations	307			307		

Notes: SD = Standard deviation. PIP = posterior inclusion probability. The frequentist check includes the variables recognized by BMA as comprising the best model. Standard errors are clustered at the study level. All variables are described in Table 4. Additional details on the BMA exercise can be found in the online appendix at `meta-analysis.cz/water`.

at all. One outcome for spatial variation that would be consistent with the previous meta-analyses would be evidence of no difference between the income elasticities for developed and developing countries. It follows that a developing country with similar structural parameters to those of a developed country can conduct similar water demand policy. This finding is close to the concept of technology adoption and supports the theory of conditional convergence. It should not, however, be applied unconditionally, as there is evidence for spatial variation across continents.

Data and publication characteristics. In accordance with Dalhuisen *et al.* (2003) and Sebri (2014), we argue that the estimates of the income elasticity of water demand are insensitive to the use of household or aggregate data. Nevertheless, the usage of daily data seems to produce systematically smaller elasticities, although BMA only suggests weak evidence for this effect. The income elasticities reported in *Published* studies are arguably smaller than those in studies coming from unrefereed sources; this effect becomes stronger with an increasing *Impact factor* of the publication outlet. It is also important to note, however, that studies reporting higher estimates attract greater attention from readers (since these papers acquire a higher number of *Citations*), but this effect is not economically significant. Thus, we identify the presence of effects unobserved by the methodological variables hidden in publication status and that the direction of their estimates is in line with the conclusions of Dalhuisen *et al.* (2003): published studies tend to report smaller elasticities than do unpublished studies.

We have established thus far that the mean estimated income elasticity of water demand, 0.27 (reported in Table 2), is influenced to a large extent by publication bias, methodology, and data heterogeneity. By accounting for publication bias in the literature we reduce the mean estimate to 0.15—when preference is also given to studies that attempt to correct for the endogeneity bias (Table 3). The estimate is, however, still not free from other potential biases resulting from data, method, and publication heterogeneity. To estimate the underlying elasticity beyond all of these effects, we construct a synthetic study that employs the preferred method, data, and publication choices and uses all of the information in the literature. Such a "best-practice" estimate is inevitably subject to the subjective decision of what the most appropriate methods,

data, and publication choices are. Therefore, we execute several robustness checks to check the sensitivity of our conclusions.

The best-practice estimate is a result of a linear combination of the BMA coefficients from Table 5 and our chosen values for the respective variables. We prefer newer studies published in outlets with a large impact factor and those receiving a high number of citations; we also prefer the use of broader data sets and methodologies that try to correct for endogeneity bias. Therefore, we set the values of the control variables of the demand equation, data with daily granularity, methods controlling for endogeneity, and publication characteristics at their sample maxima. Further, we set the values of the variables indicating the presence of publication bias, higher than daily granularity data, cross-sectional and time series data, and the estimation techniques not controlling for endogeneity at their sample minima. We leave the rest of the variables at their sample means but distinguish between the average pricing scheme and the marginal pricing scheme, as there is no clear preference for either of these schemes in the literature.

The best-practice estimation implies an elasticity of 0.082 with a 95% confidence interval of $(-0.242, 0.407)$ for the average pricing scheme and an elasticity of 0.169 with a 95% confidence interval of $(-0.155, 0.493)$ for the marginal pricing scheme. The confidence intervals are approximate and constructed using the standard errors estimated by OLS. Although the confidence intervals are wide, the plausible changes in the definition of the best practice (such as setting *Lagged variable* at the sample minimum or *Discrete-continuous* model choice at the sample maximum) changes the best-practice estimates at the third decimal place only. We conclude that the income elasticity of water demand is on average 0.15 or less and does not exceed the value of 0.5 with 95% probability, meaning that it is highly unlikely that a one-percentage-point increase in income would lead to more than a 0.5% increase in the demand for water.

In Table 6 we perform a related exercise to illustrate the sensitivity of implied elasticities to the configuration of different variables that capture the data and methods used in primary studies. We select five representative studies in the literature based on the number of citations and the impact factor of the journal where the study was published. For each study we take the configuration of data and methods preferred by the author of the particular study. Given the configuration, we compute the implied income elasticity based on our BMA exercise (note

that the implied elasticity might differ considerably from the actual elasticity reported in the paper). We hold the publication characteristics constant at the values defined earlier in our best practice exercise, so that changes in implied elasticities represent solely the differences in the configuration of data and method variables across the five studies.

Table 6: Illustrative examples of the implied effect of different configurations of study design

	Implied elasticity	95% conf. int.	
Billings & Agthe (1980)	0.170	-0.154	0.495
Hewitt & Hanemann (1995)	0.195	-0.129	0.519
Nauges & Thomas (2003)	0.410	0.086	0.734
Olmstead (2009)	0.099	-0.226	0.423
Olmstead *et al.* (2007)	0.081	-0.243	0.405

Notes: The table shows estimates of income elasticities implied by our Bayesian model averaging exercise for several prominent studies in the literature with different configurations of study design (see the main text for more details). The confidence intervals are approximate and constructed using the standard errors estimated by OLS.

The first two studies, Billings & Agthe (1980) and Hewitt & Hanemann (1995), both use monthly US data on marginal prices with an increasing tariff structure, consider lump-sum transfers, and account for evaporation in their demand functions to estimate the short-run effect. The differences between the two studies are the following: in contrast to Hewitt & Hanemann (1995), Billings & Agthe (1980) use simple OLS (not the discrete-continuous approach), do not account for rainfall in their demand function, and use time-series data (not panel data). Taken together, however, the three differences only account for a difference in the implied elasticity of 0.025. A similarly small change in the implied elasticity is generated by the differences between Olmstead (2009) and Olmstead *et al.* (2007): both studies use daily US data on marginal prices with an increasing tariff structure and account for household size, temperature, and evaporation in their demand functions. In contrast to Olmstead *et al.* (2007), Olmstead (2009) uses a panel technique estimating the short-run effect (not the discrete-continuous approach estimating the long-run effect) and accounts for rainfall in his demand function, which altogether creates a difference in the implied elasticity of merely 0.018. The study of Nauges & Thomas (2003) deviates from the previously discussed studies to a larger extent: they use French annual data on other than marginal and average prices with a flat tariff rate and do not account for climate variables, which creates a difference in in the implied elasticity of more than 0.2 in comparison to the other four studies highlighted in this exercise.

5 Robustness Checks

In this section we present additional analysis that, we believe, corroborates the results of the baseline model discussed in the previous section. Table 7 shows three specifications: study fixed effects, a subsample of older studies, and a subsample of newer studies. The fixed-effects specification is compelling, because in this way we can account for unobserved differences between studies. We have already used this approach in the examination of publication bias without additional variables. Fixed effects can be used in meta-analysis even in the examination of heterogeneity, but we have to drop many variables because some do not vary within studies (or their variation is too limited). Moreover, some studies only report a single estimate, which means they have to be eliminated from any fixed-effects analysis, too. In consequence, the amount of variation that we can explore diminishes significantly, and the specification constitutes a strict robustness check of our main results.

Indeed, the specification shows that merely three variables have a posterior inclusion probability above 0.3: standard error (a proxy for the magnitude of publication bias), the interaction of standard error and control for price endogeneity (a proxy for differences in publication bias between IV and OLS frameworks), and the use of flat tariff. The flat tariff variable was not found important in the baseline specification, but estimates computed under increasing and decreasing tariffs were found to be substantially smaller than the reference estimates. Given that the reference estimate is computed using unspecified tariffs, the finding of flat tariffs yielding substantially larger estimates compared to the reference case is consistent with the previous results and reflects that these variables are correlated (tariffs cannot be both flat and decreasing at the same time, for example).

The remaining two specifications in Table 7 are constructed by splitting the sample of studies into halves according to the average year of data they use for their analysis. This way we achieve more homogeneity within each specification, but also substantially reduce the number of degrees of freedom available for estimation. For example, the 31 "newer" studies only provide 138 estimates of income elasticities. Both specifications show evidence of publication bias that is both statistically strong (as evidenced by a high posterior inclusion probability for the variable standard error) and economically important (as evidenced by the relatively large point estimate on the variable standard error, even though here it seems that publication bias

Table 7: Study fixed effects and division by study age

Response variable:	Fixed effects			Older datasets			Newer datasets		
Income elasticity	Post. mean	Post. SD	PIP	Post. mean	Post. SD	PIP	Post. mean	Post. SD	PIP
Constant	0.000	NA	1.000	0.123	NA	1.000	0.102	NA	1.000
Standard error	0.614	0.116	1.000	0.397	0.085	1.000	2.475	0.305	1.000
SE x No endog. ctrl	-0.089	0.154	0.313	-0.016	0.066	0.095	-0.001	0.084	0.036
Water demand specification									
Household size				-0.001	0.024	0.071	-0.077	0.099	0.452
Population density	-0.012	0.083	0.064	0.156	0.176	0.519	-0.881	0.202	1.000
Temperature	-0.001	0.025	0.052	-0.037	0.079	0.236	0.140	0.114	0.692
Rainfall	0.003	0.032	0.056	0.023	0.063	0.170	-0.001	0.042	0.089
Evaporation				0.528	0.160	0.983	-0.012	0.055	0.098
Difference variable	0.000	0.081	0.051	-0.362	0.208	0.853	0.002	0.021	0.053
Lagged variable	-0.017	0.104	0.070	-0.008	0.048	0.083	0.069	0.114	0.335
Discrete-continuous	0.001	0.040	0.051	0.004	0.048	0.061	0.043	0.092	0.239
Price specification									
Marginal price	0.008	0.029	0.107	0.024	0.050	0.244	-0.005	0.028	0.078
Other price	-0.004	0.048	0.056	0.210	0.195	0.626	-0.006	0.032	0.072
Data characteristics									
Long-run elasticity	0.026	0.076	0.148	-0.007	0.033	0.099	-0.001	0.020	0.052
Household data				0.040	0.088	0.233	0.004	0.027	0.069
Daily data				-0.017	0.087	0.129	-0.085	0.142	0.354
Monthly data				0.371	0.124	0.989	-0.002	0.022	0.056
Annual data				0.131	0.231	0.324	-0.004	0.033	0.064
Cross-section	0.022	0.097	0.092	-0.140	0.135	0.612	0.002	0.036	0.086
Time series				-0.269	0.193	0.741			
Estimation technique									
No endog. ctrl	0.006	0.026	0.098	0.001	0.015	0.046	0.004	0.021	0.065
Panel technique	0.001	0.014	0.055	0.008	0.032	0.099	-0.001	0.021	0.057
Other estimator	-0.004	0.029	0.061	-0.010	0.044	0.097	-0.161	0.134	0.692
Tariff structure									
Flat tariff	0.350	0.272	0.707	0.011	0.061	0.099	0.179	0.146	0.690
Increasing tariff	-0.004	0.047	0.065	-0.121	0.120	0.621	0.002	0.020	0.055
Decreasing tariff				-0.011	0.067	0.070	-0.004	0.039	0.046
Countries examined									
Europe				-0.091	0.184	0.263	-0.047	0.087	0.291
Other location				-0.017	0.070	0.121	0.014	0.050	0.128
Developed countries							-0.009	0.040	0.103
Publication characteristics									
Publication year				0.000	0.001	0.069	0.004	0.008	0.230
Citations				0.000	0.002	0.062	0.003	0.007	0.232
Impact factor				-0.566	0.148	0.997	-0.035	0.184	0.103
Published				-0.009	0.060	0.111	-0.225	0.122	0.872
Studies	62			31			31		
Observations	307			169			138		

Notes: SD = standard deviation. PIP = posterior inclusion probability. Fixed effects = model accounting for within-study heterogeneity. Older datasets = restricted to estimates from studies using data older than the median of our sample. Newer datasets = restricted to estimates from studies using data newer than the median of our sample. All models are estimated using Bayesian model averaging employing priors suggested by Eicher *et al.* (2011).

is much stronger in newer studies). The posterior inclusion probabilities for the interaction term, however, are very small, and the expected sign of the estimated coefficients in both cases provides little consolation: the coefficients are essentially zero.

Concerning older studies, our results show that control for evaporation is especially important and typically brings much higher estimates of income elasticities. Specifying the variables in differences, in contrast, yield systematically smaller estimates. Moreover, frequency of data and the presence of time dimension in the data (as opposed to pure cross-sectional estimation) matters for the reported results. The significance of the tariff structure and the impact factor of the journal where the study is published is consistent with the results reported in the baseline specification. Concerning newer studies, we find that controlling for population density systematically affects the estimates of income elasticity. In addition, control for temperature and publication status of the study (working paper or journal) is important, which is line with the baseline model presented in the previous section.

Table 8 consists of three specifications: the first one includes 6 additional variables, the second one also considers estimates that had to be recomputed to represent constant elasticities (that is, estimates resulting from other than double-log specifications), and the third one combines the elements of the previous two specifications. Concerning the additional variables, we first consider whether the estimates preferred by the authors of primary studies are systematically different from the rest of the estimates. In meta-analysis it is often disregarded that the authors of primary studies themselves do not place the same weight on each of the estimates they report. Meta-analysts typically believe that all differences in results should be explained by "objective" variation in data, methods, and publication characteristics, but it can easily be argued that some aspects of these characteristics are difficult to observe or at least codify into a meta-analysis. (Our best-practice exercise in the previous section presents an approach that gives distinct weights to distinct estimates based on their "objective" differences in data, method, and publication characteristics.) One reason that meta-analysts typically do not collect data on preferred estimates is the inherent difficulty and subjectivity of such an analysis. While some authors make their preference explicit, others refer to several specifications or talk about the general pattern of their results. We do our best to identify one preferred estimate per study.

Table 8: Additional variables and estimates of non-constant elasticity

Response variable:	Heterogeneity			Non-constant			Heterog. & Non-const.		
Income elasticity	Post. mean	Post. SD	PIP	Post. mean	Post. SD	PIP	Post. mean	Post. SD	PIP
Constant	0.013	NA	1.000	0.137	NA	1.000	-0.155	NA	1.000
Standard error	0.948	0.127	1.000	1.143	0.107	1.000	1.133	0.106	1.000
SE x No endog. ctrl	-0.550	0.175	0.991	-1.066	0.116	1.000	-1.055	0.116	1.000
Water demand specification									
Household size	-0.001	0.010	0.024	-0.069	0.068	0.583	-0.072	0.070	0.585
Population density	0.000	0.010	0.016	-0.007	0.034	0.075	-0.009	0.039	0.080
Temperature	0.115	0.077	0.778	0.074	0.069	0.616	0.068	0.069	0.573
Rainfall	0.010	0.035	0.104	-0.006	0.029	0.077	-0.007	0.031	0.071
Evaporation	0.023	0.071	0.125	0.002	0.026	0.049	0.000	0.023	0.038
Difference variable	0.096	0.095	0.578	0.268	0.058	1.000	0.270	0.056	1.000
Lagged variable	0.055	0.095	0.300	0.001	0.011	0.025	0.000	0.010	0.019
Discrete-continuous	0.004	0.033	0.030	-0.007	0.033	0.068	-0.007	0.033	0.062
Price specification									
Marginal price	0.003	0.018	0.042	0.000	0.008	0.025	0.000	0.007	0.017
Other price	0.136	0.094	0.761	0.184	0.076	0.934	0.179	0.077	0.924
Data characteristics									
Long-run elasticity	-0.005	0.023	0.061	-0.020	0.046	0.206	-0.013	0.038	0.137
Household data	0.024	0.056	0.186	-0.001	0.010	0.033	0.000	0.009	0.024
Daily data	-0.121	0.124	0.568	-0.029	0.061	0.238	-0.031	0.063	0.242
Monthly data	0.004	0.021	0.050	0.004	0.019	0.062	0.002	0.014	0.039
Annual data	0.004	0.026	0.048	0.004	0.021	0.059	0.002	0.015	0.035
Cross-section	-0.001	0.009	0.022	-0.013	0.036	0.149	-0.016	0.041	0.172
Time series	-0.003	0.028	0.026	-0.261	0.130	0.886	-0.255	0.133	0.873
Estimation technique									
No endog. ctrl	0.022	0.046	0.229	0.179	0.048	0.996	0.179	0.047	0.998
Panel technique	0.000	0.009	0.017	0.003	0.018	0.050	0.003	0.017	0.043
Other estimator	-0.008	0.036	0.067	-0.008	0.034	0.083	-0.008	0.033	0.078
Tariff structure									
Flat tariff	-0.003	0.021	0.028	0.001	0.015	0.028	0.001	0.013	0.020
Increasing tariff	-0.071	0.077	0.531	-0.033	0.054	0.325	-0.024	0.049	0.236
Decreasing tariff	-0.128	0.176	0.407	0.000	0.016	0.024	0.000	0.014	0.017
Countries examined									
Europe	-0.004	0.023	0.047	-0.001	0.013	0.030	0.000	0.010	0.022
Other location	-0.074	0.089	0.480	-0.023	0.048	0.236	-0.011	0.035	0.118
Developed countries	0.004	0.040	0.070	0.010	0.033	0.120	0.001	0.035	0.074
Publication characteristics									
Publication year	0.000	0.001	0.024	0.000	0.001	0.067	0.000	0.001	0.045
Citations	0.006	0.006	0.553	0.012	0.006	0.902	0.011	0.006	0.850
Impact factor	-0.233	0.189	0.687	-0.144	0.178	0.464	-0.166	0.188	0.509
Published	-0.122	0.086	0.759	-0.092	0.093	0.572	-0.098	0.093	0.606
Additional variables									
Non-constant				0.098	0.099	0.566	0.084	0.097	0.493
Preferred estimate	-0.001	0.010	0.025				0.000	0.007	0.018
Self-reported income	0.017	0.044	0.164				0.001	0.012	0.028
Regional rainfall	0.025	0.033	0.420				0.008	0.019	0.188
Regional temperature	0.006	0.023	0.093				0.003	0.016	0.068
Regional income	0.011	0.024	0.197				0.026	0.033	0.474
Studies	62			72			72		
Observations	307			355			355		

Notes: SD = standard deviation. PIP = posterior inclusion probability. Heterogeneity = the benchmark model with additional explanatory variables capturing if the author prefers the estimate, if the survey respondents reported the income themselves, the climate of the researched region, and GDP per capita. Non-constant = the benchmark model estimated on the enlarged data-sample including elasticities recalculated from other than log-linear functional form. All models are estimated using Bayesian model averaging employing priors suggested by Eicher *et al.* (2011).

Next, we try to introduce more context related to the region for which the elasticity is estimated. Regions vary in terms of rainfall, temperature, and income. The authors of primary studies can (and should) control for the differences in these characteristics within their sample, but the within-sample variation is typically small. In contrast, meta-analysis enables us to compare extremely arid and wet regions (as well as rich and poor regions) for which some estimates have been reported in the literature. The primary data source for these characteristics is the particular study reporting the income elasticity; if the study does not report information on rainfall, temperature, and income, we turn to the World Bank Climate Change Knowledge Portal and IMF data on GDP (recomputed to 2016 USD according to purchasing power parity, but always corresponding to the particular year and data set investigated in the primary study).

Finally, we include a dummy variable *self-reported income*, which equals 1 if data on income are self-reported, and 0 if income data are "hard;" that is, if they are taken from census data of a statistical office or from tax records (average household's income is typically used, usually stratified per water distribution company, municipality, urban area, or community). While the census data often represent an aggregation or approximation, the survey data are problematic as well. First, some surveys report actual levels of income but often not all the questionnaires return information on income; in such cases researchers approximate the missing observations (for example, Cheesman *et al.*, 2008, replace missing observations by their subgroup average, while Mylopoulos *et al.* (2004) estimate the linear model of income dependent on household characteristics to generate the missing values). Horn (2011) notes that respondents are often likely to underestimate their real income and points to the importance of the design of the questionnaire. Second, some surveys only report income ranges, such as the one used in the Miyawaki *et al.* (2011) study of Japan prefectures.

The specification with additional variables shows results that are very similar to those of our baseline model, and none of the new variables shows posterior inclusion probability that would signify even a weak effect according to the guidelines of Jeffreys (1961). The second specification of Table 8 also includes estimates from studies that assume non-constant income elasticity of water demand. In order to become comparable to the rest of our estimates, such values have to be recomputed to elasticities evaluated at sample means. We were able to find 10 additional studies that provide 48 estimates, bringing the total number of studies and estimates in our

analysis to 72 and 355, respectively. The results show that studies assuming non-constant elasticity tend to report mean elasticities that are larger than the rest of the sample, although the posterior inclusion probability of this variable only shows a weak statistical effect. The second specification of Table 8 presents quite similar results for the remaining variables, with a couple of exceptions: First, control for household size now appears to be important (again, though, with a weak posterior inclusion probability). Second, with the extended sample of studies we find evidence for time series methods bringing systematically smaller estimates of income elasticities. Third, failure to control for price endogeneity yields even more exaggerated estimates of income elasticities than we reported previously. The last specification of the table confirms the insignificance of additional variables, presence of publication bias, and the importance of the control for price endogeneity.

6 Concluding Remarks

This paper quantitatively surveys the available estimates of the income elasticity of water demand while concentrating on three main issues not addressed by the previous meta-analyses on the topic. First, we take a closer look at publication selection bias stemming from the expected and theory-supported preferences of researchers, referees, and editors for positive and statistically significant results. Second, we focus on the problem of endogeneity bias and investigate the differences in estimates produced by different estimation techniques, still accounting for the potential influences of publication bias. Third, we investigate other sources of heterogeneity behind the estimates proposed by the previous meta-analyses of Dalhuisen *et al.* (2003) and Sebri (2014); extending their analysis, in this paper we account for the model uncertainty inherent in meta-analysis, which is due to the large number of explanatory variables, using Bayesian model averaging. We produce a best-practice estimate which suggests that the income elasticity is on average much smaller than commonly thought. This can be perceived as good news for the future availability of drinking water.

Information on the magnitude of the income elasticity of water demand is important for evidence-based policy making in both developing and developed countries. In developing countries with rapidly rising populations and per-capita income, the elasticity is crucial for projecting future water consumption and thus the design of investment into related facilities. Consequently,

a mean reported elasticity of 0.3 (as suggested by previous meta-analyses) has quantitative implications that are dramatically different from our preferred estimate of 0.15. In developed countries the need for projections of future water consumption is perhaps less acute, but still important. In warm and arid regions of developed countries, water can also function as a luxury good (used for family swimming pools and extensive lawn watering), which has consequences for water pricing policies. Moreover, water supply and demand have recently been incorporated into some integrated assessment models of climate change effects (e.g., Hejazi *et al.*, 2014), and for these, income elasticities are important, too.

The literature on the income elasticity of water demand suffers from two major problems: endogeneity bias and publication bias. Our results suggest that the estimation methods ignoring endogenous prices typically exaggerate the mean income elasticity, despite that they do not suffer from publication bias. By contrast, the estimation methods controlling for endogeneity do not suffer from endogeneity bias, but since they typically report large standard errors, publication bias causes the reported estimates computed using these methods to also exaggerate the underlying elasticity (the simple average of estimates is 0.26, while the underlying effect corrected for publication bias is 0.15). Therefore, although more consistent estimation techniques eliminate endogeneity bias at the micro level, they lead to the publication selection problem that affects the entire literature by pushing the average reported income elasticity upwards. It is difficult to disentangle these two biases without deploying the statistical tools of meta-analysis.

Our results concerning publication bias are robust to controlling for various other sources of heterogeneity at the level of estimates or studies. Apart from the aspects related to publication and endogeneity bias, several method and data characteristics wield a systematic influence on the size of the reported elasticities. Including a control for temperature in demand equations has a particularly strong impact on the estimated elasticities, as does the usage of other than marginal or average price, both factors that increase the reported elasticities. Lower data granularity and non-flat tariff systems are associated with smaller elasticities. Nevertheless, although we collect 38 aspects of study design and control for publication bias, we are still unable to explain almost 50% of the variation in the reported elasticities, which leaves ample scope for further research on this important topic.

References

Agthe, D. E. & R. B. Billings (1980): "Dynamic models of residential water demand." *Water Resources Research* **16(3)**: pp. 476–480.

Al-Najjar, F. O., E. K. Al-Karablieh, & A. Salman (2011): "Residential water demand elasticity in greater Amman area." *Jordan Journal of Agricultural Sciences* **7(1)**: pp. 93–103.

Al-Qunaibet, M. H. & R. S. Johnston (1985): "Municipal demand for water in Kuwait: Methodological issues and empirical results." *Water Resources Research* **21(4)**: pp. 433–438.

Arbues, F., M. A. Garcia-Valinas, & R. Martinez-Espineira (2003): "Estimation of residential water demand: A state-of-the-art review." *The Journal of Socio-Economics* **32(1)**: pp. 81–102.

Arbues, F., I. Villanua, & R. Barberan (2010): "Household size and residential water demand: An empirical approach." *The Australian Journal of Agricultural and Resource Economics* **54**: pp. 61–80.

Asci, S. & T. Borisova (2014): "The effect of price and non-price conservation programs on residential water demand." *2014 Annual Meeting, July 27-29, 2014, Minneapolis, Minnesota 170687*, Agricultural and Applied Economics Association.

Ashenfelter, O., C. Harmon, & H. Oosterbeek (1999): "A review of estimates of the schooling/earnings relationship, with tests for publication bias." *Labour Economics* **6(4)**: pp. 453–470.

Ayadi, M., J. Krishnakumar, & M.-S. Matoussi (2002): "A panel data analysis of residential water demand in presence of nonlinear progressive tariffs." *Working Paper 6/2012*, Department of Econometrics, University of Geneva.

Bartczak, A., A. Kopanska, & J. Raczka (2009): "Residential water demand in a transition economy: Evidence from Poland." *Water Science and Technology: Water Supply* **9(5)**: pp. 509–516.

Basani, M., J. Isham, & B. Reilly (2008): "The determinants of water connection and water consumption: Empirical evidence from a Cambodian household survey." *World Development* **36(5)**: pp. 953–968.

Billings, R. B. (1982): "Specification of block rate price variables in demand models." *Land Economics* **58(3)**: pp. 386–394.

Billings, R. B. & D. E. Agthe (1980): "Price elasticities for water: A case of increasing block rates." *Land Economics* **56(1)**: pp. 73–84.

Binet, M.-E., F. Carlevaro, & M. Paul (2014): "Estimation of residential water demand with imperfect price perception." *Environmental and Resource Economics* **59(4)**: pp. 561–581.

Binet, M.-E., F. Carlevaro, M. Paul *et al.* (2012): "Estimation of residential water demand with imprecise price perception." *Economics Working Paper 201233*, Center for Research in Economics and Management (CREM), University of Rennes 1, University of Caen and CNRS.

Carter, D. W. & J. W. Milon (2005): "Price knowledge in household demand for utility services." *Land Economics* **81(2)**: pp. 265–283.

Cheesman, J., J. Bennett, & T. V. H. Son (2008): "Estimating household water demand using revealed and contingent behaviors: Evidence from Vietnam." *Water Resources Research* **44(11)**: pp. 859–863.

Dalhuisen, J. M., R. J. Florax, H. L. de Groot, & P. Nijkamp (2001): "Price and Income Elasticities of Residential Water Demand: Why Empirical Estimates Differ." *Tinbergen Institute Discussion Papers 01-057/3*, Tinbergen Institute.

Dalhuisen, J. M., R. J. Florax, H. L. de Groot, & P. Nijkamp (2003): "Price and income elasticities of residential water demand: A meta-analysis." *Land Economics* **79(2)**: pp. 292–308.

Dalmas, L. & A. Reynaud (2004): "Residential water demand in Slovak Republic." In P. Koundouri (editor), "Econometrics Informing Natural Resources Management: Selected Empirical Analyses," volume 11, p. 83. Edward Elgar Publishing.

Darr, P., S. L. Feldman, & C. S. Kamen (1975): "Socioeconomic factors affecting domestic water demand in Israel." *Water Resources Research* **11(6)**: pp. 805–809.

Dharmaratna, D. & J. Parasnis (2011): "Price responsiveness of residential, industrial and commercial water demand in Sri Lanka." *The Asia Pacific Journal of Economics & Business* **15(2)**: p. 17.

Doucouliagos, C. & T. D. Stanley (2013): "Are all economic facts greatly exaggerated? Theory competition and selectivity." *Journal of Economic Surveys* **27(2)**: pp. 316–339.

Egger, M., G. D. Smith, M. Schneider, & C. Minder (1997): "Bias in meta-analysis detected by a simple, graphical test." *British Medical Journal* **315(7109)**: pp. 629–634.

Eicher, T. S., C. Papageorgiou, & A. E. Raftery (2011): "Default priors and predictive performance in Bayesian model averaging, with application to growth determinants." *Journal of Applied Econometrics* **26(1)**: pp. 30–55.

Espey, M., J. Espey, & W. D. Shaw (1997): "Price elasticity of residential demand for water: A meta-analysis." *Water Resources Research* **33(6)**: pp. 1369–1374.

Feldkircher, M. & S. Zeugner (2012): "The impact of data revisions on the robustness of growth determinants—a note on 'determinants of economic growth: Will data tell?" *Journal of Applied Econo-*

metrics **27(4)**: pp. 686–694.

Fenrick, S. A. & L. Getachew (2012): "Estimation of the effects of price and billing frequency on household water demand using a panel of Wisconsin municipalities." *Applied Economics Letters* **19(14)**: pp. 1373–1380.

Foster, H. S. & B. R. Beattie (1979): "Urban residential demand for water in the United States." *Land Economics* **55(1)**: pp. 43–58.

Foster, H. S. & B. R. Beattie (1981): "On the specification of price in studies of consumer demand under block price scheduling." *Land Economics* **57(4)**: pp. 624–629.

Frondel, M. & M. Messner (2008): "Price perception and residential water demand: Evidence from a German household panel." In "16th Annual Conference of the European Association of Environmental and Resource Economists in Gothenburg," pp. 25–28.

Galbraith, R. F. (1990): "The radial plot: graphical assessment of spread in ages." *International Journal of Radiation Applications and Instrumentation. Part D. Nuclear Tracks and Radiation Measurements* **17(3)**: pp. 207–214.

Garcia, S. & A. Reynaud (2004): "Estimating the benefits of efficient water pricing in France." *Resource and Energy Economics* **26(1)**: pp. 1–25.

Gaudin, S. (2005): "Using water bills to reinforce price signals: Evidence from the USA." *Water Science and Technology: Water Supply* **5(6)**: pp. 163–171.

Gaudin, S. (2006): "Effect of price information on residential water demand." *Applied Economics* **38(4)**: pp. 383–393.

Gaudin, S., R. C. Griffin, & R. C. Sickles (2001): "Demand specification for municipal water management: evaluation of the Stone-Geary form." *Land Economics* **77(3)**: pp. 399–422.

Gibbs, K. C. (1978): "Price variable in residential water demand models." *Water Resources Research* **14(1)**: pp. 15–18.

Hanemann, W. M. & C. Nauges (2005): "Heterogeneous responses to water conservation programs: The case of residential users in Los Angeles." *CUDARE Working Paper Series 1026*, University of California at Berkeley, Department of Agricultural and Resource Economics and Policy.

Hausman, J. (1978): "Specification Tests in Econometrics." *Econometrica* **46(6)**: pp. 1251–71.

Havranek, T. (2010): "Rose effect and the euro: is the magic gone?" *Review of World Economics* **146(2)**: pp. 241–261.

Havranek, T. (2015): "Measuring Intertemporal Substitution: The Importance of Method Choices and Selective Reporting." *Journal of the European Economic Association* **13(6)**: pp. 1180–1204.

Havranek, T., D. Herman, & Z. Irsova (2018): "Does Daylight Saving Save Electricity? A Meta-Analysis." *Energy Journal* **39(2)**: pp. 35–61.

Havranek, T., R. Horvath, Z. Irsova, & M. Rusnak (2015a): "Cross-country heterogeneity in intertemporal substitution." *Journal of International Economics* **96(1)**: pp. 100–118.

Havranek, T., R. Horvath, & A. Zeynalov (2016): "Natural Resources and Economic Growth: A Meta-Analysis." *World Development* **88**: pp. 134 – 151.

Havranek, T. & Z. Irsova (2011): "Estimating Vertical Spillovers from FDI: Why Results Vary and What the True Effect Is." *Journal of International Economics* **85(2)**: pp. 234–244.

Havranek, T. & Z. Irsova (2012): "Survey Article: Publication Bias in the Literature on Foreign Direct Investment Spillovers." *Journal of Development Studies* **48(10)**: pp. 1375–1396.

Havranek, T. & Z. Irsova (2017): "Do Borders Really Slash Trade? A Meta-Analysis." *IMF Economic Review* **65(2)**: pp. 365–396.

Havranek, T., Z. Irsova, & K. Janda (2012): "Demand for gasoline is more price-inelastic than commonly thought." *Energy Economics* **34(1)**: pp. 201–207.

Havranek, T., Z. Irsova, K. Janda, & D. Zilberman (2015b): "Selective reporting and the social cost of carbon." *Energy Economics* **51(C)**: pp. 394–406.

Havranek, T. & O. Kokes (2015): "Income elasticity of gasoline demand: A meta-analysis." *Energy Economics* **47(C)**: pp. 77–86.

Havranek, T., M. Rusnak, & A. Sokolova (2017): "Habit formation in consumption: A meta-analysis." *European Economic Review* **95(C)**: pp. 142–167.

Hejazi, M., J. Edmonds, L. Clarke, P. Kyle, E. Davies, V. Chaturvedi, M. Wise, P. Patel, J. Eom, K. Calvin, R. Moss, & S. Kim (2014): "Long-term global water projections using six socioeconomic scenarios in an integrated assessment modeling framework." *Technological Forecasting and Social Change* **81**: pp. 205 – 226.

Hewitt, J. A. (1993): *Watering households: The two-error discrete-continous choice model of residential water demand.* Dissertation, University of California, Berkeley.

Hewitt, J. A. & W. M. Hanemann (1995): "A discrete/continuous choice approach to residential water demand under block rate pricing." *Land Economics* **71(2)**: pp. 173–192.

Hoffmann, M., A. Worthington, & H. Higgs (2006): "Urban water demand with fixed volumetric charging in a large municipality: The case of Brisbane, Australia." *Australian Journal of Agricultural and Resource Economics* **50(3)**: pp. 347–359.

Hoglund, L. (1999): "Household demand for water in Sweden with implications of a potential tax on water

use." *Water Resources Research* **35(12)**: pp. 3853–3863.

Horn, T. (2011): "Welfare effects of access to water service in Cambodia." *Economics Bulletin* **31(3)**: pp. 2075–2089.

Hussain, I., S. Thrikawala, & R. Barker (2002): "Economic analysis of residential, commercial, and industrial uses of water in Sri Lanka." *Water International* **27(2)**: pp. 183–193.

Irsova, Z. & T. Havranek (2013): "Determinants of Horizontal Spillovers from FDI: Evidence from a Large Meta-Analysis." *World Development* **42(C)**: pp. 1–15.

Ito, K. (2013): "How Do Consumers Respond to Nonlinear Pricing? Evidence from Household Water Demand." *Workng paper*, Stanford University.

Jeffreys, H. (1961): *Theory of probability.* Oxford Classic Texts in the Physical Sciences. Oxford University Press, Oxford, 3 edition.

Jia, L. & C. Bao (2014): "Residential fresh water demand in China: A panel data analysis." *ZEF Working Paper 126*, Center for Development Research, University of Bonn.

Jones, C. V. & J. R. Morris (1984): "Instrumental price estimates and residential water demand." *Water Resources Research* **20(2)**: pp. 197–202.

Lyman, R. A. (1992): "Peak and off-peak residential water demand." *Water Resources Research* **28(9)**: pp. 2159–2167.

Mansur, E. T. & S. M. Olmstead (2012): "The value of scarce water: Measuring the inefficiency of municipal regulations." *Journal of Urban Economics* **71(3)**: pp. 332–346.

Miaou, S. (1990): "A class of time-series urban water demand models nonlinear climatic effects." *Water Resources Research* **26(2)**: pp. 169–178.

Miyawaki, K., Y. Omori, & A. Hibiki (2011): "Panel data analysis of Japanese residential water demand using a discrete/continuous choice approach." *Japanese Economic Review* **62(3)**: pp. 365–386.

Monteiro, H. & C. Roseta-Palma (2011): "Pricing for scarcity? An efficiency analysis of increasing block tariffs." *Water Resources Research* **47(6)**.

Musolesi, A. & M. Nosvelli (2007): "Dynamics of residential water consumption in a panel of Italian municipalities." *Applied Economics Letters* **14(6)**: pp. 441–444.

Mylopoulos, Y. A., A. K. Mentes, & I. Theodossiou (2004): "Modeling residential water demand using household data: A cubic approach." *Water International* **29(1)**: pp. 105–113.

Nauges, C. & J. Strand (2007): "Estimation of non-tap water demand in Central American cities." *Resource and Energy Economics* **29(3)**: pp. 165–182.

Nauges, C. & A. Thomas (2003): "Long-run study of residential water consumption." *Environmental and Resource Economics* **26(1)**: pp. 25–43.

Nauges, C. & C. Van Den Berg (2009): "Demand for piped and non-piped water supply services: Evidence from southwest Sri Lanka." *Environmental and Resource Economics* **42(4)**: pp. 535–549.

Nieswiadomy, M. L. (1992): "Estimating urban residential water demand: Effects of price structure, conservation, and education." *Water Resources Research* **28(3)**: pp. 609–615.

Nieswiadomy, M. L. & S. L. Cobb (1993): "Impact of pricing structure selectivity on urban water demand." *Contemporary Economic Policy* **11(3)**: pp. 101–113.

Nieswiadomy, M. L. & D. J. Molina (1989): "Comparing residential water estimates under decreasing and increasing block rates using household data." *Land Economics* **65(3)**: pp. 280–289.

Nieswiadomy, M. L. & D. J. Molina (1991): "A note on price perception in water demand models." *Land Economics* **67(3)**: pp. 352–359.

Olmstead, S. M. (2009): "Reduced-form versus structural models of water demand under nonlinear prices." *Journal of Business & Economic Statistics* **27(1)**: pp. 84–94.

Olmstead, S. M., W. M. Hanemann, & R. N. Stavins (2007): "Water demand under alternative price structures." *Journal of Environmental Economics and Management* **54(2)**: pp. 181–198.

Piper, S. (2003): "Impact of water quality on municipal water price and residential water demand and implications for water supply benefits." *Water Resources Research* **39(5)**.

Polycarpou, A. & T. Zachariadis (2013): "An econometric analysis of residential water demand in Cyprus." *Water Resources Management* **27(1)**: pp. 309–317.

Reynaud, A., S. Renzetti, & M. Villeneuve (2005): "Residential water demand with endogenous pricing: The Canadian case." *Water Resources Research* **41(11)**.

Rietveld, P., J. Rouwendal, & B. Zwart (1997): "Estimating Water Demand in Urban Indonesia: A Maximum Likelihood Approach to block Rate Pricing Data." *Discussion Paper 97-072/3*, Tinbergen Institute.

Saleth, R. & A. Dinar (1997): "Satisfying Urban Thirst. Water Supply Augmentation and Pricing Policy in Hyderabad City, India." *Papers 395*, World Bank - Technical Papers.

Schefter, J. E. & E. L. David (1985): "Estimating residential water demand under multi-part tariffs using aggregate data." *Land Economics* **61(3)**: pp. 272–280.

Schleich, J. & T. Hillenbrand (2009): "Determi-

nants of residential water demand in Germany." *Ecological Economics* **68(6)**: pp. 1756–1769.

Sebri, M. (2013): "Intergovernorate disparities in residential water demand in Tunisia: a discrete/continuous choice approach." *Journal of Environmental Planning and Management* **56(8)**: pp. 1192–1211.

Sebri, M. (2014): "A meta-analysis of residential water demand studies." *Environment, Development and Sustainability* **16(3)**: pp. 499–520.

Shin, J.-S. (1985): "Perception of price when price information is costly: Evidence from residential electricity demand." *The Review of Economics and Statistics* pp. 591–598.

Stanley, T. D. (2005): "Beyond Publication Bias." *Journal of Economic Surveys* **19(3)**: pp. 309–345.

Statzu, V. & E. Strazzera (2009): "Water demand for residential uses in a mediterranean region: Econometric analysis and policy implications." *Working paper*, University of Cagliari.

Strand, J. & I. Walker (2005): "Water markets and demand in Central American cities." *Environment and Development Economics* **10(3)**: pp. 313–335.

Strong, A. & V. K. Smith (2010): "Reconsidering the economics of demand analysis with kinked budget constraints." *Land Economics* **86(1)**: pp. 173–190.

Tabieh, M., A. Salman, E. Al-Karablieh, H. Al-Qudah, & H. Al-Khatib (2012): "The residential water demand function in Amman-Zarka basin in Jordan." *Wulfenia Journal* **19(11)**: pp. 324–333.

Taylor, R. G., J. R. McKean, & R. A. Young (2004): "Alternate price specifications for estimating residential water demand with fixed fees." *Land Economics* **80(3)**: pp. 463–475.

WEF (2015): "Global Risks 2015 Report." *Technical report*, World Economic Forum.

Wichman, C. J. (2014): "Perceived price in residential water demand: Evidence from a natural experiment." *Journal of Economic Behavior & Organization* **107(PA)**: pp. 308–323.

Wichman, C. J., L. O. Tylor, & R. H. von Haefen (2016): "Conservation Policies: Who Responds to Price and Who Responds to Prescription?" *Journal of Environmental Economics and Management* **79**: pp. 114–134.

Williams, M. (1985): "Estimating urban residential demand for water under alternative price measures." *Journal of Urban Economics* **18(2)**: pp. 213–225.

Williams, M. & B. Suh (1986): "The demand for urban water by customer class." *Applied Economics* **18(12)**: pp. 1275–1289.

Wong, S. (1972): "A model on municipal water demand: A case study of northeastern Illinois." *Land Economics* **48(1)**: pp. 34–44.

Yoo, S.-H. (2007): "Estimation of household tap water demand function with correction for sample selection bias." *Applied Economics Letters* **14(14)**: pp. 1079–1082.

Younes, B. Z. & M. S. Matoussi (2011): "Residential water demand: A panel cointegration approach and application to Tunisia." *Working Paper 656*, The Economic Research Forum (ERF).

Zapata, O. (2015): "More water please, it's getting hot! The effect of climate on residential water demand." *Water Economics and Policy* **1(3)**: p. 1550007.

Zigraiova, D. & T. Havranek (2016): "Bank competition and financial stability: Much ado about nothing?" *Journal of Economic Surveys* **30(5)**: pp. 944–981.