# Valuing water purification services of forests: a production function approach using panel data from China's Sichuan province

Zhaoyang Liu[1*] and Heqing Huang[2]

[1] Zhaoyang Liu, Department of Land Economy, University of Cambridge, Cambridge, UK and [2] National School of Development, Peking University, Beijing, China

*Corresponding author. E-mail: zl290@cam.ac.uk

## ONLINE APPENDIX

## Appendix A

**Table A1.** Previous econometric studies on the value of forests' water purification services

| Reference | Study area | Full sample size | Panel data | Catchment delineation | Estimation method | Main results |
|---|---|---|---|---|---|---|
| Abildtrup *et al*. (2013) | Vosges, France | 232 water supply services | No | No | Water prices regressed against forestland; spatial econometric model | One point increase in the proportion of forestland (agricultural land decreased by the same extent) leads to a €0.015/$m^3$ decrease in the price of drinking water. |
| Abildtrup *et al*. (2015) | Vosges, France | 232 water supply services | No | No | Spatial switching regression model | Spatially lagged forest cover significantly reduces the price of drinking water. |
| Fiquepron *et al*. (2013) | France | 93 administrative departments | No | No | Simultaneous equations model | One point increase in the proportion of forests would imply a €0.0034/$m^3$ decrease in the price of drinking water. |
| Lopes *et al*. (2019) | Portugal | 235 water treatment firms | No | No | Water treatment costs regressed against forests; 2SLS model where the volume of treated water was instrumented | Elasticity of treatment costs w.r.t. forest cover is –0.0564% for ground water firms and 0.0217% (statistically insignificant) for surface water firms. |
| Piaggio and Siikamäki (2021) | Costa Rica | 20 water treatment plants, monthly data, 1,315 obs. for turbidity, 1,156 obs. for aluminium sulphate | Yes | Yes | Two step approach (turbidity regressed against forests, usage of aluminium sulphate regressed against turbidity); fixed effects panel data model | Elasticity of the usage of aluminium sulphate w.r.t. avoided conversion of forests: –0.026%. |
| Singh and Mishra (2014) | Greater Mumbai, India | Water quality: 6 sites, monthly data, Jan. 1998–Dec. 2010; treatment costs: 1 water treatment plant, | Partly | Yes | Two step approach (turbidity regressed against forests, treatment costs regressed against turbidity); mixed effects | 1% decrease in forest cover will increase treatment costs by 1.58%. |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | monthly data, Apr. 1995 – Mar. 2011 | | | panel data model; time series model | |
| Vincent *et al*. (2016) | Perak, Malaysia | 41 water treatment plants, monthly data, 3,894 obs. | Yes | Yes | Water treatment costs regressed against forests; fixed effects panel data model | Elasticities of treatment costs w.r.t. virgin and logged forests are –0.47% and –0.31%, respectively. |
| Vincent *et al*. (2020) | Thailand | 162 Provincial Waterworks Authority branches, quarterly data, 5,843 obs. | Yes | Yes | Material costs regressed against land uses (forests excluded as the reference land use); fixed effects panel data model | Elasticity of material costs w.r.t. deforestation ranges from 0.25 to 0.57 (depending on whether forests are converted to agriculture, urban, or miscellaneous land use). |
| Westling *et al*. (2020) | Sweden | Water quality: 76 water treatment plants (WTP), monthly data, 7,981 obs. for *E.coli*, 4,076 obs. for turbidity; chemical costs: 20 WTPs, annual data, 248 obs. | Yes | Yes | Two step approach (water quality regressed against forests, chemical costs regressed against water quality); dynamic panel data model | 1% decrease in forest cover would give a 0.228% increase in chemical costs. |

*Notes:* The studies by Ernst (2004), Freeman *et al*. (2008), Cunha *et al*. (2016) and Warziniack *et al*. (2017) analysed small datasets and/or did not control for variables other than land use, and are therefore less comparable with the more formal econometric studies mentioned above. Knowler *et al*. (2017) quantitatively simulated how water treatment costs are affected by the presence of timber harvesting and the usage of forest roads, but did not provide a direct valuation of the water purification services of forests. Some other studies have looked at how water treatment costs are affected by land use types other than forests (such as cropland and ranchland, as in McDonald *et al*. (2016), and agricultural versus non-agricultural land, as in Forster and Murray (2007)), or the impact of forest cover on water quality rather than on water treatment costs (such as O'Donoghue *et al*. (2021) and Segurado *et al*. (2018)). These studies notably differ from the formal econometric studies that we build upon.

**Table A2.** Definition and description of variables

| | Mean | SD | 5% quantile | 95% quantile |
|---|---|---|---|---|
| Panel 1: Variables in the water treatment cost function (county level, obs. = 1,618) | | | | |
| Unit water treatment cost (CNY/$m^3$) | 1.07 | 0.65 | 0.25 | 2.34 |
| Forestland (%), inside catchment, 0–3 km | 18.82 | 25.04 | 0 | 77.31 |
| Forestland (%), outside catchment, 0–3 km | 14.71 | 19.50 | 0 | 58.44 |
| Cropland (%), inside catchment, 0–3 km | 43.69 | 29.91 | 0.55 | 91.24 |
| Urban areas (%), inside catchment, 0–3 km | 5.49 | 11.44 | 0 | 26.11 |
| Water supply (1 mn $m^3$/yr.) | 15.83 | 50.55 | 0.84 | 41.93 |
| Wage rate (CNY 1k/yr.) | 23.79 | 19.51 | 3.11 | 61.71 |
| Pct. of private water supply firms | 27.88 | 28.80 | 0.00 | 80.00 |
| Rainfall 0–3 km (mm/yr.) | 984.46 | 148.80 | 720.64 | 1214.00 |
| Panel 2: Variables in the Heckman correction (obs. = 4,080) | | | | |
| Ethnic minority autonomous county (binary: 0 = no; 1 = yes) | 0.30 | 0.46 | 0 | 1 |
| Distance to Chengdu (km) | 223.86 | 126.00 | 57.78 | 474.82 |
| Urban population (1k people) | 87.09 | 84.30 | 5.30 | 242.00 |
| Number of phones per capita | 0.31 | 0.38 | $4.29\times10^{-3}$ | 1.03 |
| Road density (km/$km^2$) | 0.65 | 0.70 | 0.06 | 1.85 |
| Percentage of private water works | 23.07 | 29.16 | 0.00 | 80.00 |

*Note:* CNY6.62 = USD1 in 2018 prices.

**Table A3.** Estimated probit sample selection equation

| **Dependent variable:** Water treatment data observed<br>**Explanatory variables:** | **Model A1** |
|---|---|
| ***Ethnic minority autonomous county*** | ***–0.51******** |
| | ***(0.06)*** |
| ***Distance to Chengdu*** | ***$-5.89\times10^{-4}$****** |
| | ***($2.06\times10^{-4}$)*** |
| ***Urban population*** | ***$1.65\times10^{-3}$****** |
| | ***($3.42\times10^{-4}$)*** |
| ***Number of phones per capita*** | ***–0.13**** |
| | ***(0.07)*** |
| ***Road density*** | ***0.06**** |
| | ***(0.04)*** |
| ***Percentage of private water works*** | ***$1.32\times10^{-3}$**** |
| | ***($7.55\times10^{-4}$)*** |
| Number of observations | 4,080 |
| Model significance (*p*-value) | 0.00 |
| McFadden's $R^2$ | 0.07 |

*Notes:* Estimates that are statistically significant in both models are highlighted in bold italics (up to the 10% significance level). Asterisks indicate statistical significance: **p*-value < 0.10, ****p*-value < 0.01. Standard errors are in parentheses.

**Table A4.** Estimated water treatment cost function controlling for forest cover at farther distances

| **Dependent variable:** IHS of unit cost | **Model A2** |
|---|---|
| **Explanatory variables:** | |
| ***IHS of pct. of forestland 0–3 km (inside catchment)*** | ***$-2.91\times10^{-2}$***** |
| | ***($1.12\times10^{-2}$)*** |
| IHS of pct. of forestland 3–6 km (inside catchment) | $3.65\times10^{-2}$ |
| | ($3.62\times10^{-2}$) |
| IHS of pct. of forestland 6–10 km (inside catchment) | $-1.86\times10^{-2}$ |
| | ($3.08\times10^{-2}$) |
| IHS of pct. of cropland 0–3 km (inside catchment) | $2.16\times10^{-2}$ |
| | ($2.31\times10^{-2}$) |
| IHS of pct. of urban area 0–3 km (inside catchment) | $1.62\times10^{-2}$ |
| | ($1.28\times10^{-2}$) |
| IHS of rainfall 0–3 km | 5.32 |
| | (8.13) |
| Squared IHS of rainfall 0–3 km | –0.22 |
| | (0.33) |
| ***IHS of water supply*** | ***–0.21****** |
| | ***(0.04)*** |
| IHS of wage rate | $1.11\times10^{-2}$ |
| | (0.08) |
| IHS of pct. of state owned water works | $5.27\times10^{-3}$ |
| | ($5.81\times10^{-3}$) |
| ***Inverse Mills Ratio*** | ***–0.79****** |
| | ***(0.29)*** |
| County fixed effects | Yes |
| Year fixed effects | Yes |
| Clustered standard errors (at the county level) | Yes |
| Number of observations | 1,618 |
| Model significance (*p*-value) | 0.00 |
| $R^2$ (within) | 0.75 |

*Notes:* Estimates that are statistically significant in both models are highlighted in bold italics (up to the 10% significance level). Asterisks indicate statistical significance: ***p*-value < 0.05, ****p*-value < 0.01. Standard errors are in parentheses.

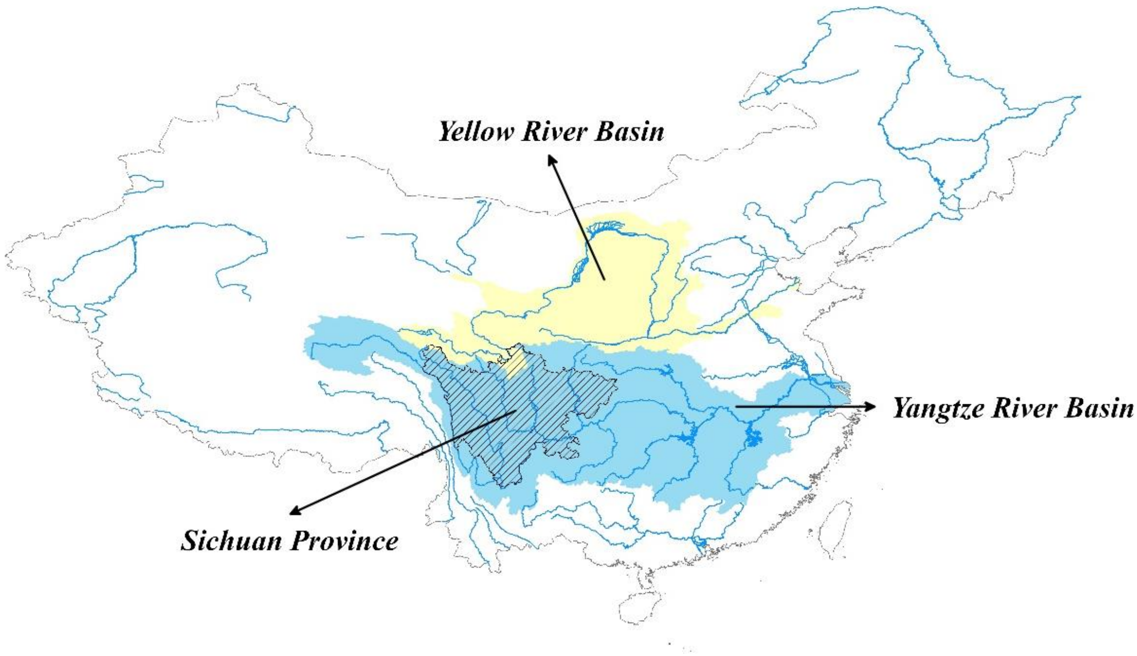
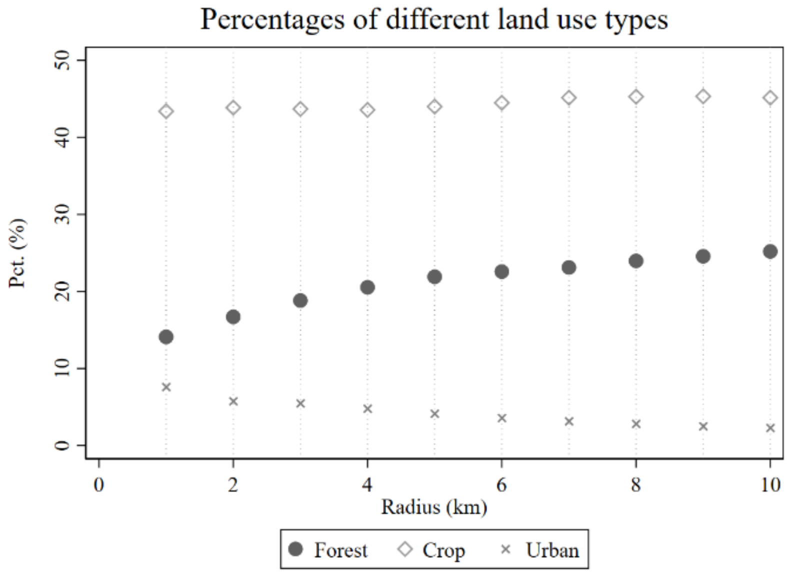



**Figure A1.** Location of Sichuan province on a watershed map of China.

*Source of the watershed map*: The Chinese Academy of Sciences.

*a*




*b*

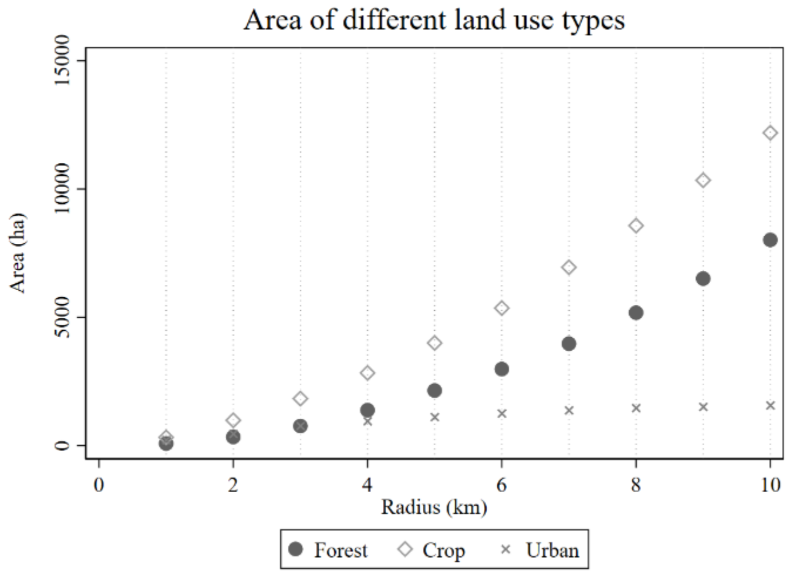



**Figure A2.** Sample mean percentages (panel ***a***) and area (panel ***b***) of different land use types.

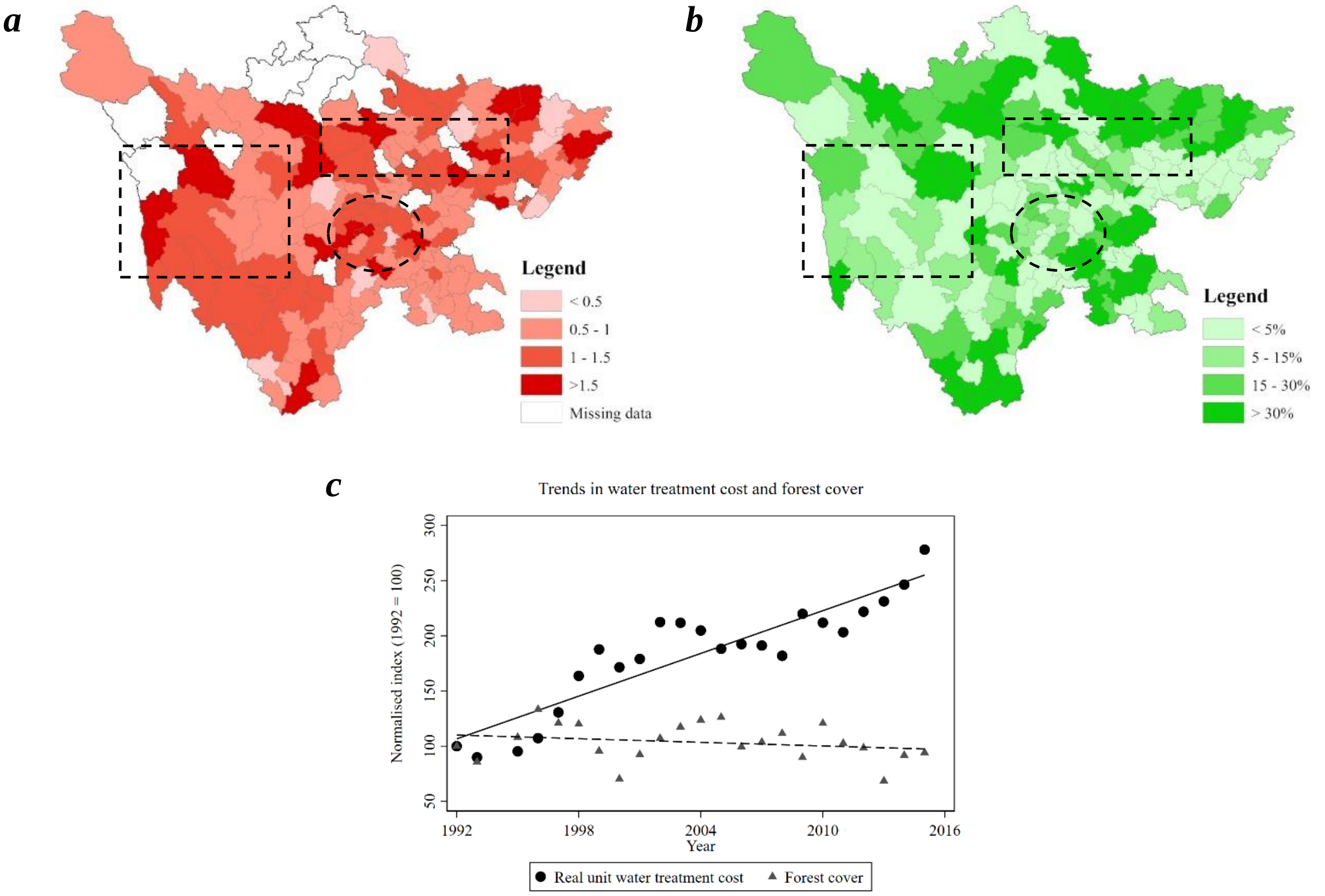
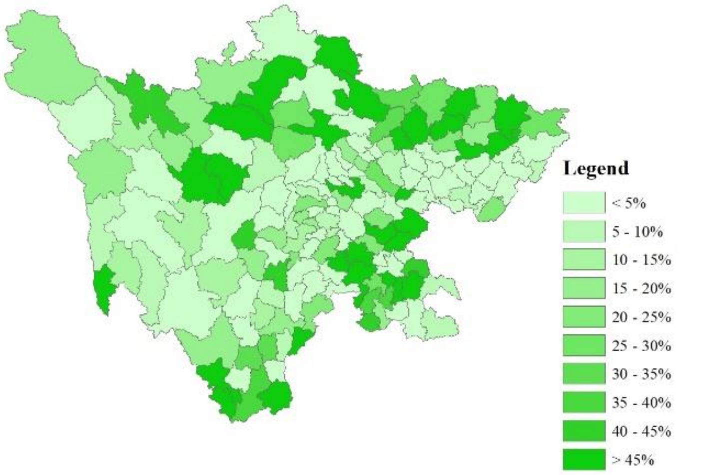



**Figure A3.** Spatial and temporal patterns of drinking water treatment cost and forest cover.

*Notes:* Panel ***a*** shows spatial distribution of the unit water treatment cost (CNY/m$^3$, CNY6.62 = USD1 in 2018 prices); Panel ***b*** shows spatial distribution of the percentage of forestland 0–3 km upstream; Panel ***c*** shows time trends in the unit water treatment cost (inflation-adjusted) and forest cover. Cut-off values in panels ***a*** and ***b*** roughly represent quantiles.

*a*




*b*

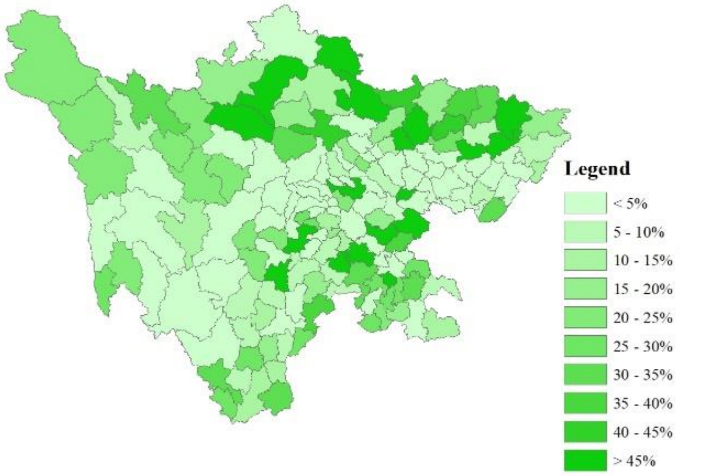



**Figure A4.** Spatial distribution of forest cover (3 km radius) inside and outside the catchment area.

*Notes:* Panel ***a*** shows percentage of forestland (3 km radius) inside the catchment; Panel ***b*** shows percentage of forestland (3 km radius) outside the catchment.

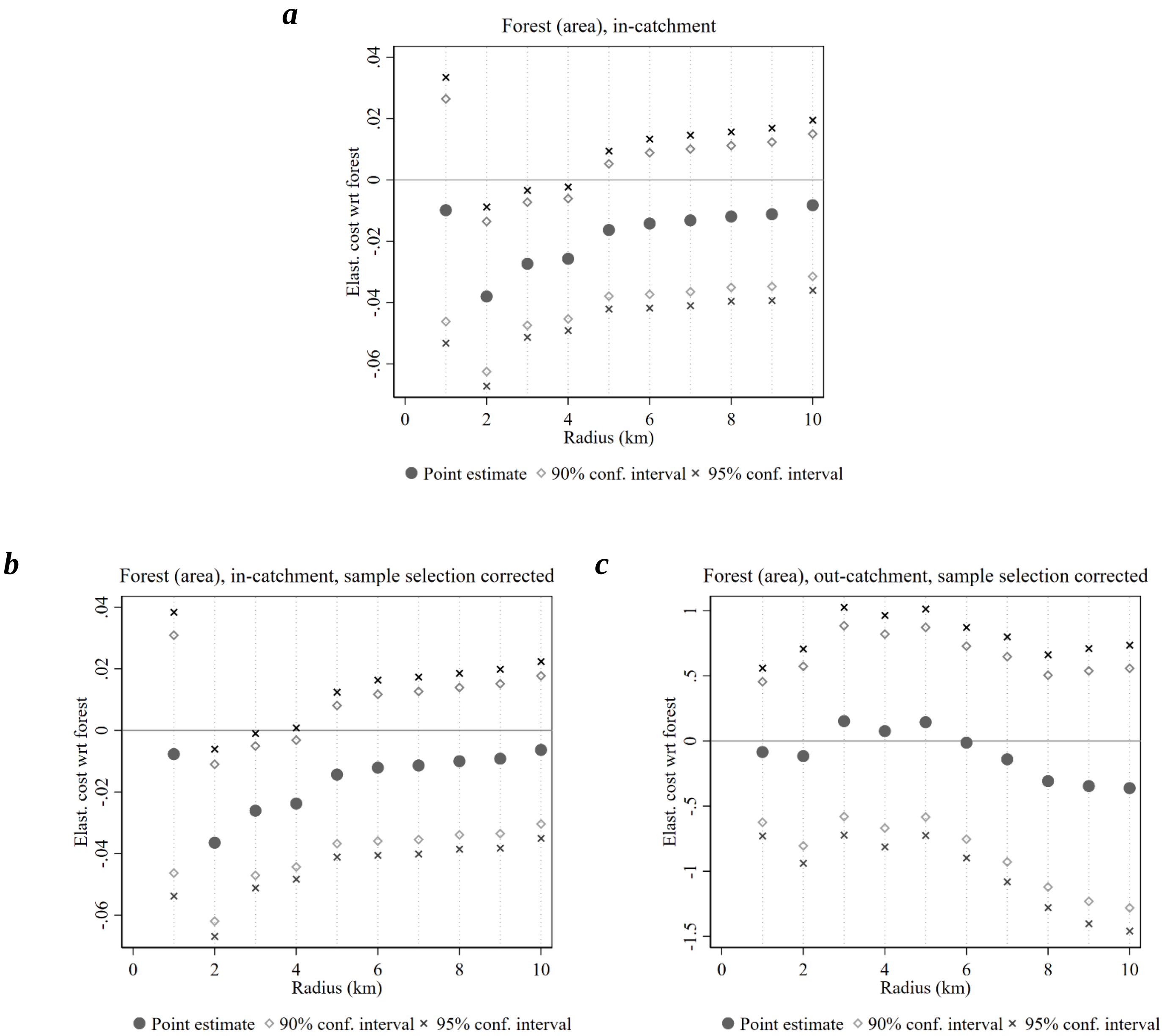



**Figure A5.** Elasticity estimates with respect to forestland *area* within different radiuses.

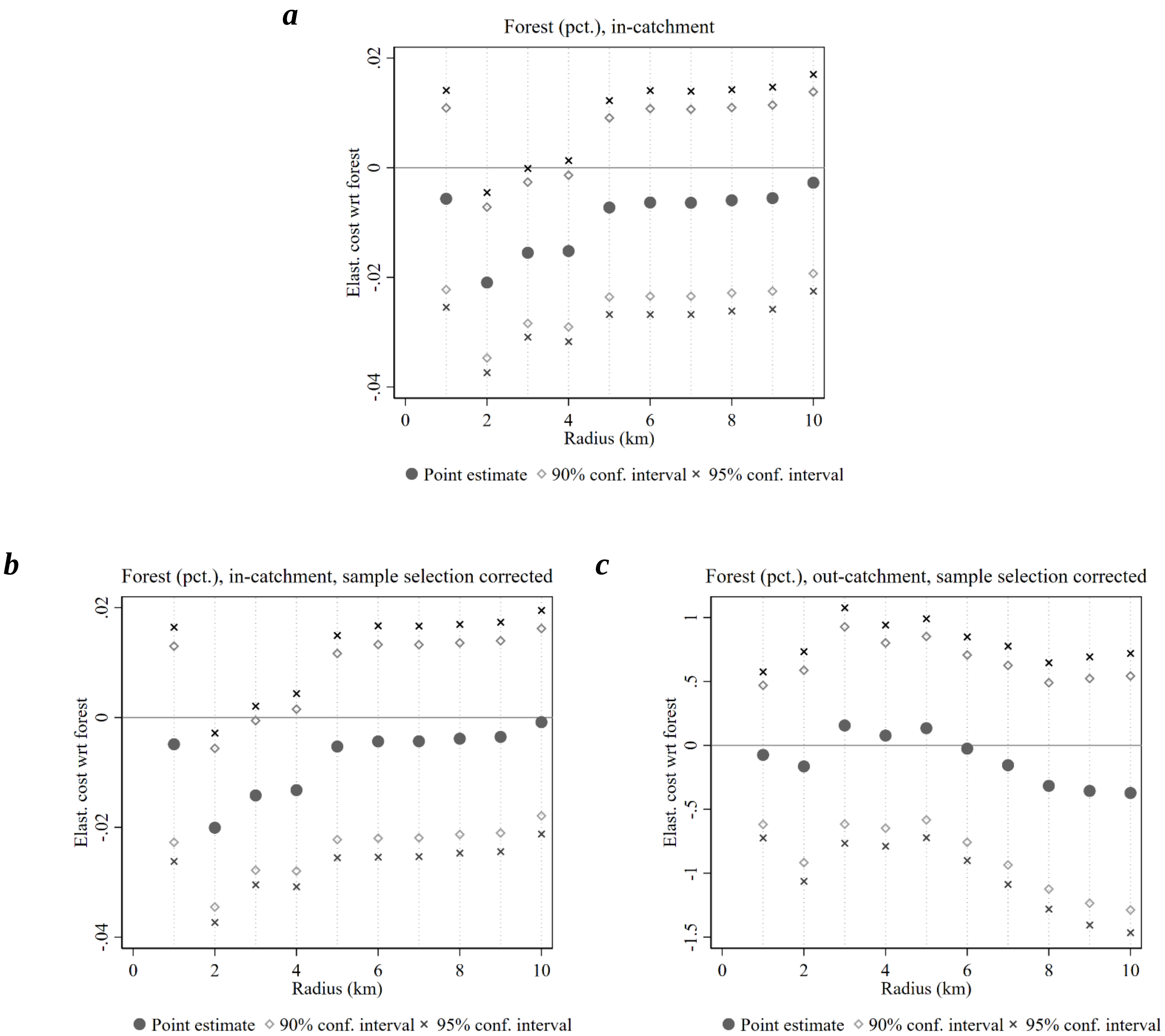



**Figure A6.** Elasticity estimates with respect to forestland percentages within different radiuses, logarithmically transformed variables**.**

**Appendix B**

We repeated the analysis using an alternative form of the production function which accounts for both the quantity and quality of output, following Grieco and McDevitt (2017):

$$zq^{\sigma} = \boldsymbol{x_1^{\alpha} x_2^{\beta} e^{\gamma}}, \tag{A1}$$

which basically assumes that producing the same amount of output *z* with higher quality *q* requires a higher level of aggregate production input.[1] It is assumed that the production process can be divided into separable and independent steps, where producers first decide the quantity and quality of output, and then in the next step decide the levels of production inputs. In that case, the cost function can be derived in the same manner as in section 2, through solving a cost minimisation problem with predetermined levels of *z* and *q*:

$$\bar{c} = \frac{c}{\tilde{z}} = \frac{w_1 x_1^* + w_2 x_2^*}{\tilde{z}} = \Omega \boldsymbol{w_1}^{\frac{\alpha}{\alpha+\beta}} \boldsymbol{w_2}^{\frac{\beta}{\alpha+\beta}} \boldsymbol{\tilde{e}}^{-\frac{\gamma}{\alpha+\beta}} \tilde{z}^{\left(\frac{1}{\alpha+\beta}-1\right)} \tilde{q}^{\frac{\sigma}{\alpha+\beta}}, \tag{A2}$$

which can be rewritten as the log linear form:

$$ln\bar{c} = ln\Omega + \frac{\boldsymbol{\alpha}}{\boldsymbol{\alpha+\beta}} ln\boldsymbol{w_1} + \frac{\boldsymbol{\beta}}{\boldsymbol{\alpha+\beta}} ln\boldsymbol{w_2} - \frac{\boldsymbol{\gamma}}{\boldsymbol{\alpha+\beta}} ln\boldsymbol{\tilde{e}} + \left(\frac{1}{\boldsymbol{\alpha+\beta}} - 1\right) ln\tilde{z} + \frac{\sigma}{\boldsymbol{\alpha+\beta}} ln\tilde{q}. \tag{A3}$$

The empirical implication is that this production function allows the quality of treated water to be explicitly controlled for in the regression models as an explanatory variable. We next re-estimated all regression models controlling for the quality of treated water, and the results are almost identical to those presented in the main text, as can be seen in figure A7.

[1] Grieco and McDevitt (2017) described the production function in the log linear form: $lnz = \mu(\boldsymbol{\alpha} ln\boldsymbol{x_1} + \boldsymbol{\beta} ln\boldsymbol{x_2} + \boldsymbol{\gamma} ln\boldsymbol{e})$, and $lnq = \frac{1-\mu}{\sigma}(\boldsymbol{\alpha} ln\boldsymbol{x_1} + \boldsymbol{\beta} ln\boldsymbol{x_2} + \boldsymbol{\gamma} ln\boldsymbol{e})$, where the parameter $\mu$ distinguishes the proportions of aggregate input used to produce the quantity and quality of output. It can be seen that the production function described in equation (A1), after taking the logarithm of both sides, would be equivalent to the model proposed by Grieco and McDevitt (2017). We expressed the production function in the original Cobb-Douglas functional form (instead of the log linear form) to facilitate the derivation of the cost function.

**Table A5.** Estimated water treatment cost function controlling for the quality of treated water

| **Dependent variable:** IHS of unit cost<br>**Explanatory variables:** | **Model B1** |
|---|---|
| ***IHS of pct. of forestland 0–3 km (inside catchment)*** | ***$-1.67\times10^{-2}$**** |
| | ***($1.00\times10^{-2}$)*** |
| IHS of pct. of cropland 0–3 km (inside catchment) | $1.85\times10^{-2}$ |
| | ($2.37\times10^{-2}$) |
| IHS of pct. of urban area 0–3 km (inside catchment) | $1.66\times10^{-2}$ |
| | ($1.22\times10^{-2}$) |
| IHS of rainfall 0–3 km | 5.43 |
| | (8.12) |
| Squared IHS of rainfall 0–3 km | –0.23 |
| | (0.33) |
| ***IHS of water supply*** | ***–0.21****** |
| | ***(0.04)*** |
| IHS of wage rate | $-6.23\times10^{-3}$ |
| | (0.08) |
| IHS of pct. of state owned water works | $5.40\times10^{-3}$ |
| | ($5.83\times10^{-3}$) |
| IHS of quality | 0.07 |
| | (1.28) |
| ***Inverse Mills Ratio*** | ***–0.78****** |
| | ***(0.29)*** |
| County fixed effects | Yes |
| Year fixed effects | Yes |
| Clustered standard errors (at the county level) | Yes |
| Number of observations | 1,618 |
| Model significance (*p*-value) | 0.00 |
| $R^2$ (within) | 0.75 |

*Notes:* Estimates that are statistically significant in both models are highlighted in bold italics (up to the 10% significance level). Asterisks indicate statistical significance: *$p$-value < 0.10, ***$p$-value < 0.01. Standard errors are in parentheses.

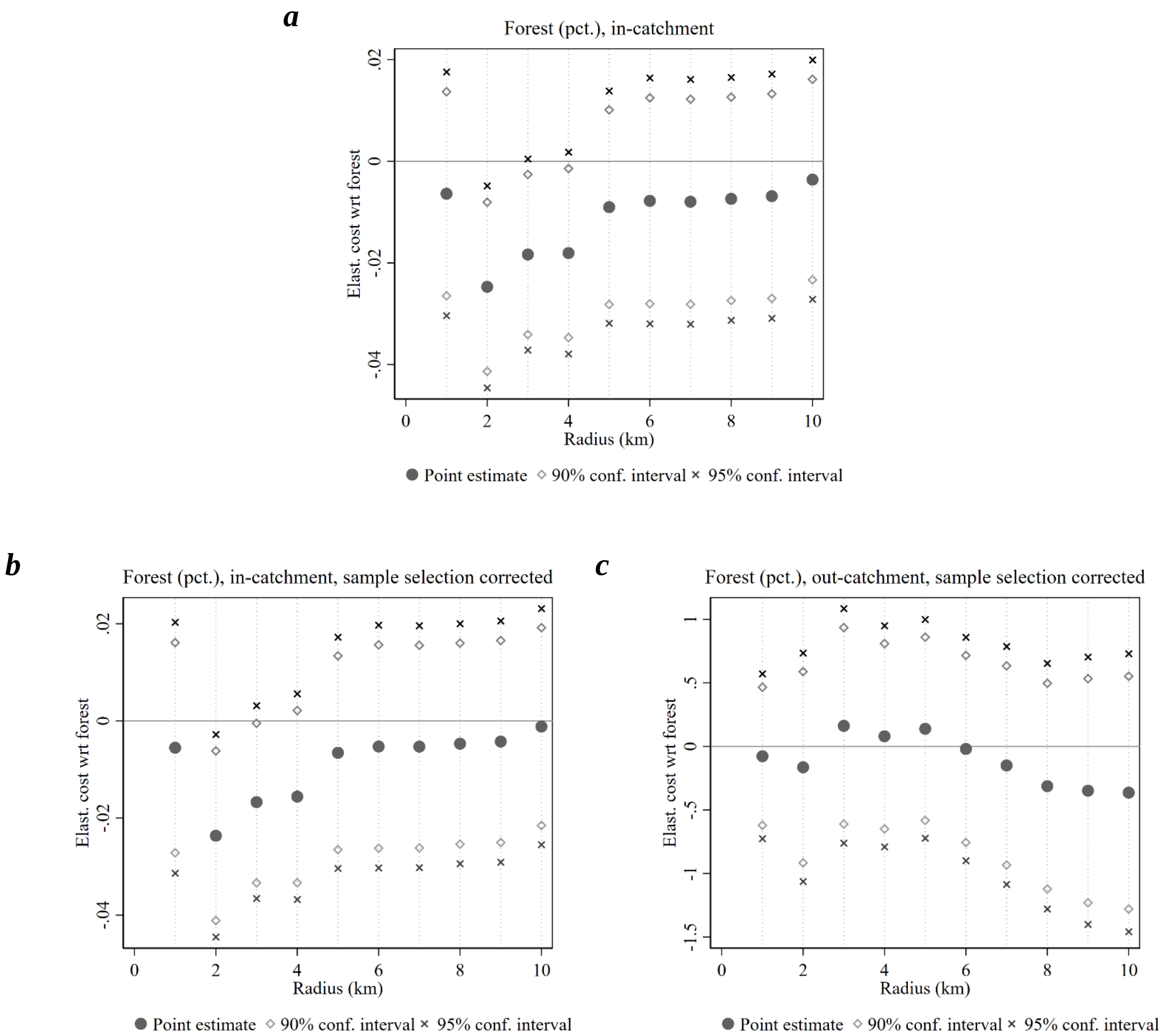



**Figure A7.** Elasticity estimates with respect to forestland percentages within different radiuses controlling for the quality of treated water.

**Appendix C**

Lastly, we explored the dependency of forests' water purification effects on rainfall. On the one hand, higher rainfall levels may increase sediment loads in surface water by washing away eroded soil and other debris into surface water (Vincent *et al*., 2016). In that case, forest cover may deliver a higher level of water purification services by intercepting rainfall, regulating streamflow and hence reducing soil erosion and sediment loads, which suggests a positive relationship between rainfall levels and forests' water purification effects. On the other hand, forest cover may have limited ability to intercept rainfall and regulate streamflow in case of prolonged periods of very high levels of rainfall (Calder *et al*., 2007), which implies that forests' water purification effects may negatively depend on rainfall levels. Both cases imply that the future scale of this ecosystem service may be affected by climate change. Figure A8 shows an increasing trend in annual rainfall in Sichuan province during the study period. If rainfall indeed affects forests' water purification services, we would be able to predict the future trajectory of the level of this ecosystem service as a function of predicted rainfall in the context of climate change. This would provide useful insights as to the benefits and costs of climate change mitigation.

We estimated Models C1 and C2 (table A6) to explore the nexus between rainfall and the implications of forest cover for drinking water treatment costs. The two models are identical to Models 1 and 2 (table 1 in the main text), except that the two new models include an interaction term between forest cover and rainfall, '*IHS. pct. forestland 0–3 km (inside catchment)* × *IHS. rainfall 0–3 km*', which captures whether rainfall influences the effect of forest cover on drinking water treatment costs. The forest and rainfall variables were centred before they were added to the models and the interaction term, so that the coefficients on the two variables can still be independently interpreted as the main effects at the mean values of other regressors. In both Models C1 and C2, it can be seen that the estimate on this interaction term is small in size (about 10 per cent of the main effect of forest cover) and statistically indistinguishable from zero ($p$-value > 0.85 in both models). Therefore, the jury is out on the hypothesised relation between rainfall and forests' water purification service. However, this is likely associated with the nature of our annual-level dataset, which has precluded us from making full use of the variation of the two variables at a higher temporal resolution. Further research is warranted in light of the importance of this research question.

**Table A6.** Dependency of forests' water purification services on rainfall

| **Dependent variable:** IHS of unit cost<br>**Explanatory variables:** | **Model C1** | **Model C2** |
|---|---|---|
| ***IHS of pct. of forestland 0–3 km*** | ***–1.83×10$^{-2}$**** | ***–1.67×10$^{-2}$**** |
| ***(inside catchment)*** | ***(9.53×10$^{-3}$)*** | ***(1.01×10$^{-2}$)*** |
| IHS of pct. of cropland 0–3 km | 1.53×10$^{-2}$ | 1.85×10$^{-2}$ |
| (inside catchment) | (2.29×10$^{-2}$) | (2.37×10$^{-2}$) |
| IHS of pct. of urban area 0–3 km | 1.72×10$^{-2}$ | 1.66×10$^{-2}$ |
| (inside catchment) | (1.22×10$^{-2}$) | (1.22×10$^{-2}$) |
| IHS of rainfall 0–3 km | –2.00×10$^{-2}$ | –7.17×10$^{-2}$ |
| | (0.10) | (0.10) |
| Squared IHS of rainfall 0–3 km | –0.23 | –0.23 |
| | (0.33) | (0.33) |
| IHS. pct. forestland 0–3 km (inside catchment) × | –2.49×10$^{-3}$ | –1.21×10$^{-3}$ |
| IHS. rainfall 0–3 km | (1.69×10$^{-2}$) | (1.68×10$^{-2}$) |
| ***IHS of water supply*** | ***–0.20**** | ***–0.21**** |
| | ***(0.04)*** | ***(0.04)*** |
| IHS of wage rate | 2.45×10$^{-3}$ | –6.55×10$^{-3}$ |
| | (0.08) | (0.08) |
| IHS of pct. of state owned water works | 6.75×10$^{-3}$ | 5.40×10$^{-3}$ |
| | (5.89×10$^{-3}$) | (5.81×10$^{-3}$) |
| ***Inverse Mills Ratio*** | | ***–0.78**** |
| | | ***(0.29)*** |
| County fixed effects | Yes | Yes |
| Year fixed effects | Yes | Yes |
| Clustered standard errors (at the county level) | Yes | Yes |
| Number of observations | 1,618 | 1,618 |
| Model significance ($p$-value) | 0.00 | 0.00 |
| $R^2$ (within) | 0.75 | 0.75 |

*Notes:* Estimates that are statistically significant in both models are highlighted in bold italics (up to the 10% significance level). Asterisks indicate statistical significance: *$p$-value < 0.10, ***$p$-value < 0.01. Standard errors are in parentheses.

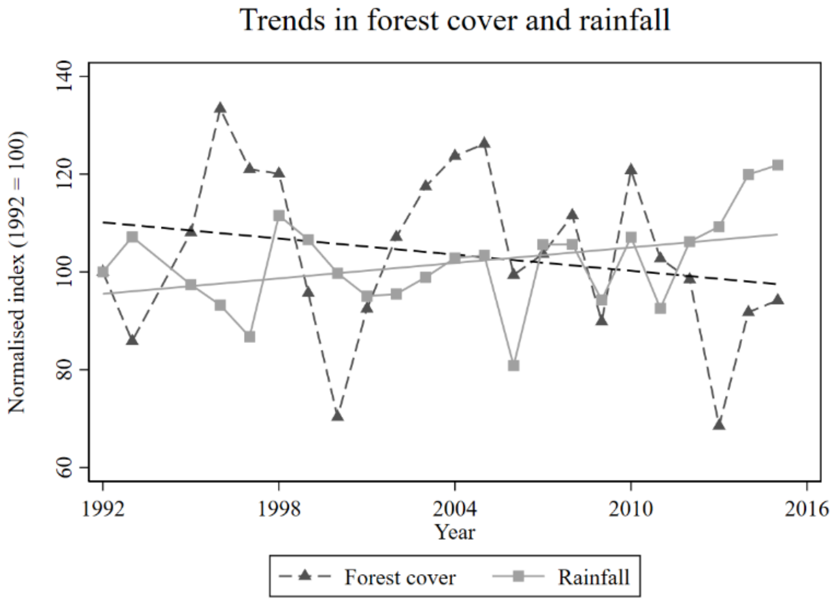



**Figure A8.** Time trends in forest cover and rainfall.

**References**

**Abildtrup J, Garcia S and Kere E** (2015) Land use and drinking water supply: a spatial switching regression model with spatial endogenous switching. *Revue d'Économie Régionale & Urbaine*, Mai(1–2), 321–342.

**Abildtrup J, Garcia S and Stenger A** (2013) The effect of forest land use on the cost of drinking water supply: a spatial econometric analysis. *Ecological Economics* **92**, 126–136.

**Calder I, Hofer T, Vermont S and Warren P** (2007) Vers une nouvelle compréhension des arbres et des forêts. *Unasylva* **58**, 3–10 (in French).

**Cunha DGF. Sabogal-Paz LP and Dodds WK** (2016) Land use influence on raw surface water quality and treatment costs for drinking supply in São Paulo State (Brazil). *Ecological Engineering* **94**, 516–524.

**Ernst C** (2004) *Protecting the Source: Land Conservation and the Future of America's Drinking Water*. San Francisco: Trust for Public Land.

**Fiquepron J, Garcia S and Stenger A** (2013) Land use impact on water quality: valuing forest services in terms of the water supply sector. *Journal of Environmental Management* **126**, 113–121.

**Forster DL and Murray C** (2007) Effects of pesticide use and farming practices on water treatment costs in Maumee River Basin communities. In Hitzhusen FJ (ed), *Economic Values of River Systems*. Norhampton, MA: Edward Elgar Publishing Limited, pp. 115–128.

**Freeman J, Madsen R and Hart K** (2008) *Statistical Analysis of Drinking Water Treatment Plant Costs, Source Water Quality, and Land Cover Characteristics*. San Francisco: Trust for Public Land.

**Grieco PLE and McDevitt RC** (2017) Productivity and quality in health care: evidence from the dialysis industry. *The Review of Economic Studies* **84**, 1071–1105.

**Knowler D, Page A, Cooper A and Araujo HA** (2016) Valuing a logging externality: loss of the water purification service of temperate coastal rainforests. *Water Economics and Policy* **3**, 1650032.

**Lopes AF, Macdonald JL, Quinteiro P, Arroja L, Carvalho-Santos C, Cunha-e-Sá MA and Dias AC** (2019) Surface vs. groundwater: the effect of forest cover on the costs of drinking water. *Water Resources and Economics* **28**, 100123.

**McDonald RI, Weber KF, Padowski J, Boucher T and Shemie D** (2016) Estimating watershed degradation over the last century and its impact on water-treatment costs for the world's large cities. *Proceedings of the National Academy of Sciences* **113**, 9117–9122.

**O'Donoghue C, Buckley C, Chyzheuskaya A, Green S, Howley P, Hynes S, Upton V and Ryan M** (2021) The spatial impact of rural economic change on river water quality. *Land Use Policy* **103**, 105322.

**Piaggio M and Siikamäki J** (2021) The value of forest water purification ecosystem services in Costa Rica. *Science of The Total Environment* **789**, 147952.

**Segurado P, Almeida C, Neves R, Ferreira MT and Branco P** (2018) Understanding multiple stressors in a Mediterranean basin: combined effects of land use, water scarcity and nutrient enrichment. *Science of The Total Environment* **624**, 1221–1233.

**Singh S and Mishra A** (2014) Deforestation-induced costs on the drinking water supplies of the Mumbai metropolitan, India. *Global Environmental Change* **27**, 73–83.

**Vincent JR, Ahmad I, Adnan N, Burwell WB, Pattanayak SK, Tan-Soo J-S and Thomas K** (2016) Valuing water purification by forests: an analysis of Malaysian panel data. *Environmental and Resource Economics* **64**, 59–80.

**Vincent JR, Nabangchang O and Shi C** (2020) Is the distribution of ecosystem service benefits pro-poor? Evidence from water purification by forests in Thailand. *Water Economics and Policy* **6**, 2050005.

**Warziniack T, Sham CH, Morgan R and Feferholtz Y** (2017) Effect of forest cover on water treatment costs. *Water Economics and Policy* **3**, 1750006.

**Westling N, Stromberg PM, and Swain RB** (2020) Can upstream ecosystems ensure safe drinking water – insights from Sweden. *Ecological Economics* **169**, 106552.