Appendices (For Online Publication Only)

M. Lipscomb and A. M. Mobarak,
"Decentralization and Pollution Spillovers:
Evidence from the Re-drawing of County Borders
in Brazil"

A History of County Splits in Brazil

Brazil has a dynamic history of county splitting dating back to 1930. The topic has been investigated and debated in the literature on public administration in Brazil, particularly during the early and mid-1990s when the country made significant constitutional changes affecting the county splitting process. The attitudes of scholars and policymakers toward the phenomenon range from dismissive to adamantly supportive. County splits have been discounted as "phantom counties" in the case of remote areas of the Amazon (Lordello de Mello, 1971) or "irresponsible emancipation" redistributing public funds away from population centers and toward excessive administrative structures in small counties (Affonso, 1996; Gomes and MacDowell, 2000). Others (notably Noronha and Cardoso 1996) defend the process as a democratic right to self-determination and a means to bring basic services and infrastructure to impoverished areas, thereby preventing "rural exodus" to large cities. While there remain questions about the efficacy of new counties, many studies have examined the causes of county splitting. The literature suggests that in addition to demographic trends such as population growth and economic expansion, the history of county splits has also been driven by prevailing political regimes governing the process and the structure of fiscal incentives offered to local governments.

A.1 Reasons for Splits

Brazilian districts have chosen to become new counties for multiple reasons. Based on a review of the literature, Cachatori and Cigolini (2013) summarize the most commonly cited factors in county splitting:

- Original counties were too extensive in territory, leaving remote areas neglected by the municipal government.
- Population growth occurred in outlying areas, especially as roads were expanded.
- Lack of basic services such as water, sanitation, electricity, schools and health services, and/or the presence of poverty prompted communities to seek independence. Lordello de Mello (1992) presents several cases studies of communities who received free pumped water, clinics or electricity only after splitting off into new counties.
- In contrast to the reasoning of impoverishment and neglect, many new districts claimed that they had sufficient economic activity and fiscal resources to be viable as an independent county.

The above reasons are often given by defenders of county splitting, and they echo the findings of Bremaeker (1996b), who surveyed the mayors of newly created counties. A parallel line of argument focuses on political and institutional factors, including corruption and financial incentives:

- County splits can be initiated by local politicians to gain political favor, and other rent-seeking behavior.
- Waves of county splitting respond to institutional attitudes toward decentralization, including federal funding regulations.

A.2 Evolution of the County Splitting Process

Table A1 shows the historical trends in county splitting in Brazil since 1940. There are two distinct waves of rapid county creation, from 1940-1963 and from 1988-1996, with a hiatus from 1964-1987 when the number of counties actually decreased. These waves were driven almost exclusively by Brazil's prevailing political system, which either restricted or encouraged county splits.

Table A1: Growth in Number of Counties, 1940-Present

Time Period	Number of Counties (Beginning)	Newly Created	Percent Growth
1940-1949	1,574	315	20.00%
1950-1959	1,889	887	46.40%
1960-1963	2,766	1,459	52.70%
1964-1969	4,235	-283	-6.70%
1970-1979	3,952	39	1.00%
1980-1989	3,991	433	10.80%
1990-1996	4,424	1,074	24.30%
1997-2012	5,498	67	1.20%
2013	5,565		

Sources: Bremaeker (1996a), Cachatori and Cigolini (2013).

Early Years of Municipal Federalism

Brazil's first wave of county splits took place in the 1940s. The Constitution of 1934 began the tradition of what Gomes and MacDowell (2000) termed "municipal federalism", mandating shared tax revenues between the federal government, states and counties. In prior years, only state industrial and professional taxes were shared with counties.

The Constitution of 1946 further encouraged county splitting by mandating a minimum entitlement to each local government: the Fundo de Participação dos Municípios (FPM). When it was first introduced, the FPM was a flat allotment to counties regardless of size, which prompted state governments to encourage county splitting as a means of attracting additional federal funds (Lordello de Mello, 1992). The legal process of county splits was determined by state government, and was generally lax. Accordingly, the number of Brazilian counties grew dramatically through the 1950s and early 1960s, until the military coup of 1964 ended civilian rule.

Military Rule, 1964-1985

Brazil was ruled by a centralized junta of military technocrats for two decades following the coup. A new constitution was instituted in 1967 which transferred decisions on many local matters to the federal government, including the sole power to create new counties. New restrictions required that new counties have a minimum population of 100,000 or no less than 5

thousandths of the state's population with no less than 5 thousandths of the state's tax revenue (Lordello de Mello, 1992). As a result, almost no county splits occurred during this period, and in the first five years of the dictatorship counties were consolidated such that the total number dropped by nearly 7 percent.

Constitution of 1988

The junta leaders oversaw a gradual transition to civilian rule, and in 1985 Brazil held its first presidential election in over 20 years. With the return to democracy, a new constitution was written which returned to a shared revenue and legal structure between the federal government and municipalities. In the tradition of pre-1964 politics, the Constitution of 1988 defined the Federal Republic of Brazil as the "indivisible union of the states, municipalities and of the Federal District" of Brasilia. As such, municipal governments were given the sole responsibility of providing public infrastructure including transportation, water, sewage and sanitation, and primary responsibility for social services (MacLachlan, 2003). States once again governed the county split process, and splits would be decided by plebiscite of voters in the area wishing to secede.

Within the first five years of the new constitution, 885 new counties were formed. During the first 13 years of decentralization (1984-1997), 1,405 new counties formed, most of them with fewer than 20,000 inhabitants, and over half of them "micro-counties" of fewer than 5,000 (Gomes and MacDowell, 2000). The move toward decentralization released pent-up demand for the creation of new counties, but there was also a fiscal incentive to split. The constitution restored the FPM, and its disproportionate tax entitlements for small counties. Of the new counties, 650 had populations of under 10,000 (MacLachlan, 2003).

Constitution Amendment of 1996

Following the post-1988 wave of decentralization, there was a political backlash aimed at putting the brakes on county splits. In 1996, the Brazilian Congress passed a constitutional amendment which placed three restrictions on the county splitting process: (1) regulations would continue to be established by state law, but within a period determined by federal law; (2) proposed new counties must pass a Municipal Viability Study prior to holding the plebiscite for secession; (3) plebiscites must involve populations of both the original county and the new county, not just the separating population Bremaeker (1996a)

While debating the amendment, defenders of county splitting (notably IBAM economists and geographers such as Bremaeker (1996a) and Noronha (1996)) argued that policymakers were overreacting to what was essentially a backlog of demand for new counties resulting from the repression during the military era. They noted that the growth rate in counties from 1988 to 1996 was still well below the rates of the 1950s and early 1960s. Bremaeker compared the backlash to the early years of the dictatorship, which also reacted strongly to a rapid wave of county splits in the early 1960s.

The amendment passed and had the desired effect, limiting the growth in counties to near one percent since it took effect. Noronha (1996) argued that the new plebiscite rules would make it less likely to reach a quorum in plebiscite elections, since residents of the original county who receive satisfactory services from local government will not be motivated to vote.

Moving Forward

State regulations on county split requirements still vary significantly across the country. Cachatori and Cigolini (2013) classify the requirements into five different types: (i) minimum population; (ii) length of existence of the district trying to secede; (iii) the district's share of state tax revenue; (iv) distance to the county seat; and (v) level of urban development. All states have a minimum population requirement, although the specific number varies by region. Cachatori and Cigolini (2013) find that legislation in the Northeast is more restrictive than in the Southeast – that is, otherwise similar districts in terms of population have a more difficult emancipation process if they are located in the Northeast.

In 2013, the Brazilian Congress was still considering legislation to impose national population minimums for new counties. Proposta de Lei Complementar 416 proposed minimums of 5,000 in the North and Central-West, 7,000 in the Northeast, and 10,000 in the South and Southeast (National Congress of Brazil, 2008). Cachatori and Cigolini (2013) conduct a simulation to predict which districts would be eligible to create new counties under the proposed federal rules versus under current state legislative regimes. Their results differ dramatically according to the specific legislation adopted. Overall, the federal proposal would result in 50 percent fewer eligible districts nationwide, although this is heavily influenced by the densely populated South and Southeast regions, where current rules are less restrictive. In the North and Northeast regions, the federal proposal would relax the requirements, resulting in 27 percent more eligible districts.

Role of the FPM

The FPM is a federal tax revenue entitlement due to each county. It is a flat amount determined by population, and often cited as an incentive for superfluous county splitting. According to MacLachlan (2003), 75 percent of Brazilian counties relied on federal and state funds for over 90 percent of their budgets. The proportion is even higher for small counties, and even the largest cities could not generate 50 percent of their budget from their own tax revenues.

The rules determining the amount of the FPM have evolved along with the county split legislation described above. Initially the FPM was a flat rate for all counties, which granted a disproportionate amount of tax resources to residents of small counties. With the revival of the FPM under decentralization, the federal transfer amount was determined by population. During the 1990s, which saw the peak of the second wave of county splits, populations lost during a split were counted twice for calculating the FPM. That is, the population of counties of origin remained the same, while the newly created county was estimated from the most recent census (Noronha, 1996). This meant that splitting counties did not lose any tax revenue as a result of secession.

B General Model

We start from the regulator's problem described in section 3 of the paper. The regulator in a county will determine where to allow emissions by balancing the added utility of upstream constituents from consumption of polluting goods with the costs on downstream constituents according to the following pointwise welfare function:

$$W(x, q_x) = f(x)(u(q_x) - cq_x) - \int_x^b q_x e^{-\alpha[t-x]} f(t)dt \tag{B.1}$$

Integrating over the area before the border:

$$W(q) = \int_0^b [f(x)[u(q_x) - cq_x] - \int_x^b q_x e^{-\alpha(t-x)} f(t)dt]dx$$

where $q = \{q_x\}_{x\in[0,b]}$ is the regulator's pollution allowance. Using the calculus of variations and maximizing over q yields the first order

$$f(x)(u'(q_x^*) - c) - \int_x^b e^{-\alpha(t-x)} f(t)dt = 0 \tag{B.2}$$

Which implies:

$$u'(q_x^*) - c = \frac{\int_x^b e^{-\alpha(t-x)} f(t)dt}{f(x)},$$

so that the net marginal benefit of additional production at x is equated with the impact on the downstream population. Note that when $x = b$, the impact on the downstream population is zero, and the regulator simply sets the marginal benefit of production equal to marginal cost.

B.1 Comparative Statics

B.1.1 Prediction 1: Emissions increase going toward the Downstream Border

Taking partial derivatives of the first order condition with respect to x and q, respectively, yield

$$\Delta_{xq} = f'(x)(u'(q_x^*) - c) + f(x) - \int_x^b \alpha e^{-\alpha(t-x)} f(t)dt$$

$$\Delta_{qq} = f(x)u''(q_x^*).$$

Applying the implicit function theorem yields

$$\frac{\partial q_x^*}{\partial x} = -\frac{\Delta_{xq}}{\Delta_{qq}} = \frac{f'(x)(u'(q_x^*) - c) + f(x) - \int_x^b \alpha e^{-\alpha(t-x)} f(t)dt}{-u''(q_x^*)f(x)} \tag{B.3}$$

note that from equation B.2 we know that:

$$u'(q_x^*) - c = \frac{\int_x^b e^{-\alpha(t-x)} f(t)dt}{f(x)}$$

substituting this into $\frac{\partial q_x^*}{\partial x}$ and taking the limit as x approaches b, we get:

$$\lim_{x \to b} \frac{(f'(x) - \alpha) \int_x^b e^{-\alpha(t-x)} f(t)dt + f(x)}{f(x)(-u''(q_x^*))} = \frac{1}{-(u''(q_b^*))} \tag{B.4}$$

Note that we know by the concavity of the utility function that $u''(q_x) < 0$, so we must have $\lim_{x \to b} \frac{\partial q_x *}{\partial x} > 0$ Therefore, allowed emissions increase as the point at which they are measured approaches the downstream boundary. (This is prediction 1).

B.1.2 Prediction 2: Emissions accelerate toward the downstream border

To show that the allowed emissions are accelerating in distance to the upstream border, we need to show that $\frac{\partial^2 q_x}{\partial x^2} > 0$. We start by taking the derivative of equation B.3:

$$\begin{aligned} &\frac{[f''(x)(u'(q_x^*) - c) + f'(x)u''(q_x^*)\frac{\partial q_x^*}{\partial x} + f'(x) + \alpha f(x) - \int_x^b \alpha^2 e^{-\alpha(t-x)} f(t)dt](-u''(q_x^*))f(x)}{(-u''(q_x^*)f(x))^2} \\ &- \frac{[-u'''(q_x^*)\frac{\partial q_x^*}{\partial x} f(x) - u''(q_x^*)f'(x)] \left(f'(x)(u'(q_x^*) - c) + f(x) - \int_x^b \alpha e^{-\alpha(t-x)} f(t)dt\right)}{(-u''(q_x^*)f(x))^2} \end{aligned} \tag{B.5}$$

The denominator is squared, so it must always be positive. By the first order condition, we know that $u'(q_x^*) = c$ as $x \to b$, so the integral terms will drop out, along with any $u'(q_x^*) - c$ terms. Then the numerator of B.5 as $x \uparrow b$ is

$$(f'(b)u''(q_b^*)\frac{\partial q_b^*}{\partial x} + f'(b) + \alpha f(b))(-u''(q_b^*))f(b) + [u'''(q_b^*)\frac{\partial q_b^*}{\partial x} f(b) - u''(q_b^*)f'(b)][f(b)]$$

The first term equals

$$\left(f'(b)\left\{u''(q_b^*)\frac{\partial q_b^*}{\partial x} + 1\right\} + \alpha f(b)\right)(-u''(q_b^*))f(b)$$

To sign this term, note first that population is empirically decreasing near the boundaries so that $f'(b) < 0$. Then it would be sufficient for the term in braces, $u''(q_b^*)\frac{\partial q_b^*}{\partial x} + 1$, to be weakly negative. But we know that in the limit $\frac{\partial q_b^*}{\partial x} = \frac{-1}{u''(q_b^*)}$, so we have

$$u''(q_b^*)\frac{-1}{u''(q_b^*)} + 1 \leq 0,$$

which implies

$$\underbrace{f'(b)}_{<0} \underbrace{\left\{u''(q_b^*)\frac{\partial q_b^*}{\partial x} + 1\right\}}_{\leq 0} \geq 0$$

Substituting this back into the first term, we have:

$$\underbrace{\Big(\underbrace{f'(b)}_{<0}\underbrace{[u''(q_b^*)\frac{\partial(q_b^*)}{\partial x} + 1]}_{\leq 0} + \underbrace{\alpha f(b)}_{>0}\Big)\underbrace{(-u''(q_b^*))f(b)}_{>0}}_{>0}$$

And the second term can be rewritten

$$-u''(q_b^*)\left[-\frac{u'''(q_b^*)}{u''(q_b^*)}\frac{\partial q_b^*}{\partial x}f(b)+f'(b)\right]f(b).$$

Then the above quantity is positive since $f'(b) < 0$ and $u'''(q_b^*) > 0$. While the assumption that $u'''(q) > 0$ is less common than concavity, $u''(q) < 0$, it is a necessary condition for households to exhibit decreasing absolute risk aversion. Then both term (1) and term (2) are weakly positive, which then implies that $\frac{\partial^2 q_x}{\partial x^2} > 0$.

B.1.3 Prediction 3: Total pollution increases by strictly less after the border crossing than before the border crossing.

Let $b_0, b_1, ..., b_k, ..., b_K$ be the sequence of borders in $[0, 1]$, so that $b_k < b_{k+\ell}$, $\ell \geq 1$. To show this prediction is true, first note that it is sufficient to show that the limit of q_x^* from the left is greater than the limit of q_x^*, from the right. This implies that there is a discontinuous drop in pollution at the boundary, so that pollution increases by strictly less after the border than before, or

$$\lim_{x\uparrow b_k}\frac{\partial q_x^*}{\partial x} > \lim_{x\downarrow b_k}\frac{\partial q_x^*}{\partial x}.$$

Note that from the left,

$$\lim_{x\uparrow b_k}\frac{\partial q_x^*}{\partial x} = \frac{1}{-u''(q_{b_k}^*)} > 0,$$

as before. Let b_k' be the solution to regulator $k+1$'s problem from the right, and taking the limit yields

$$\lim_{x\downarrow b_k}\frac{\partial q_x^*}{\partial x} = \frac{(f'(b_k)-\alpha)\int_{b_k}^{b_{k+1}} e^{-\alpha(t-x)}f(t)dt + f(b_k)}{-f(b_k)u''(q_x^*)},$$

or

$$\underbrace{\frac{(f'(b_k)-\alpha)\int_{b_k}^{b_{k+1}} e^{-\alpha(t-x)}f(t)dt}{-f(b_k)u''(q_{b_k}')}}_{<0}\underbrace{-\frac{1}{u''(q_{b_k}')}}_{>0}.$$

Note that because the first term is negative, a sufficient condition for $\lim_{x\uparrow b_k}\frac{\partial q_x^*}{\partial x} > \lim_{x\downarrow b_k}\frac{\partial q_x^*}{\partial x}$ is that $\frac{-1}{u''(q*_{b_k^*})} > \frac{-1}{u''(q_{b_k}')}$. This is true if $u''(q_{b_k}') < u''(q*_{b_k})$. We also know that $q_{b_k}' < q*_{b_k}$ by the first order condition and $\partial q_x^*/\partial x > 0$, and since $u''' > 0$ is assumed, it follows that $u''(q_{b_k}') < u''(q*_{b_k})$. This implies the prediction.

B.1.4 Prediction 4: Pollution is increasing in the number of borders crossed

To show this prediction is true, we will prove the following. First, eliminating a border so that the regulator controls a larger territory leads him to allow less pollution. Second, this has no impact on the behavior of the other regulators. Consequently, adding borders leads to smaller territories and more pollution overall.

Consider removing the border at b_{l+1} in the figure below:

b_l ——————————— b_{l+1} ——————————— b_{l+2}

Regulator l is then responsible for the entire segment $[b_l, b_{l+2}]$. Notice that on $[b_{l+1}, b_{l+2}]$, q_x — the allowed pollution level — doesn't change. This is because the first order condition for all such x — $f(x)(u'(q_x*) - c) - \int_x^{b_{l+2}} e^{-\alpha(t-x)} f(t)dt$ — doesn't depend on b_{l+1}. We now need to show that if b_{l+1} increases, then q_x decreases on $[b_l, b_{l+1}]$, ie, $\frac{\partial q_x}{\partial b_{l+1}} < 0$ for $x \in [b_l, b_{l+1}]$. Starting with the first-order condition,

$$f(x)(u'(q_x*) - c) - \int_x^{b_{l+1}} e^{-\alpha(t-x)} f(t)dt = 0$$

and totally differentiating with respect to b_{l+1} yields

$$f(x)u''(q_x*)\frac{\partial q_x*}{\partial b_{l+1}} - e^{-\alpha(b_{l+1}-x)} f(b_{l+1}) = 0$$

Therefore, we can solve for $\frac{\partial q_x*}{\partial b_{l+1}}$ to see that:

$$\frac{\partial q_x*}{\partial b_{l+1}} = e^{\alpha(b_{l+1}-x)} \frac{f(b_{l+1})}{f(x)} \frac{1}{u''(q_x*)} < 0$$

Overall, this implies

$$\int_{b_l}^{b_{l+2}} q_x dx < \int_{b_l}^{b_{l+1}} q_x dx + \int_{b_{l+1}}^{b_{l+2}} q_x dx,$$

where the left-hand side is total pollution before removing the boundary, and the right-hand side is the total pollution afterwards generated by regulator l. Now note that removing this border fails to change the behavior of the other regulators. The reason is that the costs of inherited pollution enter each regulator's objective function linearly[1], so that each regulator k only internalizes the consequences of his pollution allowances on the population in his district. Consequently, pollution is increasing in the total number of borders crossed.

C Solutions under several alternative specifications

This appendix provides solutions to the optimization problem under various assumptions about population distributions, functional forms of downstream costs, and welfare functions. For computational simplicity, rather than including marginal costs of emissions for the emitter, the welfare function used in this section includes a constraint on total per person (or firm) emissions ($q_x \leq \bar{q}$). Section C1 outlines our theory under three different population distributions, Section C2 outlines our theory using non-atomistic point source polluters, Section C3 outlines our theory using a convex pollution cost function, and Section C4 outlines our theory under the assumption that the upstream county partially weights the costs they impose on the downstream county after a split.

[1]Linear pollution costs are consistent with the way pollution is commonly modeled and measured in the environmental economics literature. For example, Arceo et al (2012) find that health costs are linear in PM-10 levels. Currie et al (2013) study the effects of water pollution on fetal health using a linear specification. Currie and Neidell (2005) study the effects of air pollution on infant health using a linear specification. Relaxing this assumption is an interesting project for future work.

C.1 Different Distributions of Population

In this section we model the pollution and emissions functions under three hypothetical population distributions.

First, we define three welfare functions: W, the welfare function prior to the county split; W^u, the welfare function for the upstream county after the county split; and W^d, the welfare function for the downstream county after the county split.

$$W = f(x) \cdot ln(q_x) - \int_x^1 q_x \cdot e^{-(t-x)} \cdot f(t) \cdot d(t) \qquad s.t.\ q_x < \bar{q} \tag{C.1}$$

$$W^u = f(x) \cdot ln(q_x^u) - \int_x^{0.5} q_x^u \cdot e^{-(t-x)} \cdot f(t) \cdot d(t) \qquad s.t.\ q_x^u < \bar{q}, \quad 0 \leq x \leq 0.5 \tag{C.2}$$

$$W^d = f(x) \cdot ln(q_x^d) - \int_x^1 q_x^d \cdot e^{-(t-x)} \cdot f(t) \cdot d(t) \qquad s.t.\ q_x^d < \bar{q}, \quad 0.5 \leq x \leq 1 \tag{C.3}$$

where $f(\cdot)$ is a population density function.

We derive the following first order conditions from the welfare functions.

FOC for (C.1):

$$f(x) \cdot \frac{1}{q_x} = \int_x^1 e^{-(t-x)} \cdot f(t) \cdot d(t) + \lambda \tag{C.4}$$

FOC for (C.2):

$$f(x) \cdot \frac{1}{q_x^u} = \int_x^{0.5} e^{-(t-x)} \cdot f(t) \cdot d(t) + \lambda \tag{C.5}$$

FOC for (C.3):

$$f(x) \cdot \frac{1}{q_x^d} = \int_x^1 e^{-(t-x)} \cdot f(t) \cdot d(t) + \lambda \tag{C.6}$$

Where λ is defined as the shadow value of our $\bar{q}$ constraint of 40 ($\bar{q} = 40$ was selected to be high enough such that it does not pay a quantitatively important role in determining the shape of the pollution function).[2]

Now, we substitute the three population density functions, which are defined in Equations (C.7), (C.10) and (C.13) and displayed in Figure C1, into our first order conditions to solve for q_x^*, the optimal per-person emissions levels for the county before the split; q_x^{u*}, the optimal per-person emissions in the upstream section of the split county (i.e. $x \in [0, 0.5]$); and q_x^{d*}, the optimal per-person emissions in the downstream section of the split county (i.e. $x \in [0.5, 1]$). Since (C.4) and (C.6) are the same equation we will obtain that $q_x^{d*} = q_x^*$ for $x \in [0.5, 1]$, thus we will only display solutions for q_x^* and q_x^{u*} throughout this appendix.

[2]The $\bar{q} = 40$ constraint is added to the model only to avoid arbitrarily large emissions near downstream exit borders. In all of our theoretical extensions, $\bar{q}$ does not play a crucial role in maximization problem.

1. Uniform Population Density

   Let

$$f(x) = \begin{cases} 1 & \text{if} \quad 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases} \tag{C.7}$$

   which implies

$$q_x^* = \min\left(\frac{1}{1 - e^{(x-1)}}, \bar{q}\right) \tag{C.8}$$

$$q_x^{u*} = \min\left(\frac{1}{1 - e^{(x-0.5)}}, \bar{q}\right) \tag{C.9}$$

2. Triangular Population Density

   Let

$$f(x) = \begin{cases} 4x & \text{if} \quad 0 \leq x \leq 0.5 \\ 4(1-x) & \text{if} \quad 0.5 < x \leq 1 \\ 0 & \text{otherwise} \end{cases} \tag{C.10}$$

   which implies

$$q_x^* = \begin{cases} \min\left(\dfrac{x}{x + e^{(x-1)} - 2e^{(x-0.5)} + 1}, \bar{q}\right) & \text{if} \quad 0 \leq x < 0.5 \\ \min\left(\dfrac{1-x}{e^{(x-1)} - x}, \bar{q}\right) & \text{if} \quad 0.5 \leq x < 1 \end{cases} \tag{C.11}$$

$$q_x^{u*} = \min\left(\frac{4x}{4x - 6e^{(x-0.5)} + 4}, \bar{q}\right) \qquad \text{if} \quad 0 \leq x < 0.5 \tag{C.12}$$

3. Triangular Bimodal Population Density

   Let

$$f(x) = \begin{cases} 8x & \text{if} \quad 0 \leq x < .25 \\ 8(0.5 - x) & \text{if} \quad .25 \leq x < 0.5 \\ 8(x - 0.5) & \text{if} \quad 5 \leq x < .75 \\ 8(1-x) & \text{if} \quad .75 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases} \tag{C.13}$$

   which implies

$$q_x^* = \begin{cases} \min\left(\dfrac{x}{x + e^{(x-1)} + 2e^{(x-0.5)} - 2e^{(x-0.25)} - 2e^{(x-0.75)} + 1}, \bar{q}\right) & \text{if} \quad 0 \leq x < 0.25 \\ \min\left(\dfrac{4 - 8x}{-8x + 8e^{(x-1)} + 16e^{(x-0.5)} - 16e^{(x-0.75)} - 4}, \bar{q}\right) & \text{if} \quad 0.25 \leq x < 0.5 \\ \min\left(\dfrac{8x - 4}{8x + 8e^{(x-1)} - 16e^{(x-0.75)} + 4}, \bar{q}\right) & \text{if} \quad 0.5 \leq x < 0.75 \\ \min\left(\dfrac{1 - x}{e^{(x-1)} - x}, \bar{q}\right) & \text{if} \quad 0.75 \leq x \leq 1 \end{cases} \tag{C.14}$$

$$q_x^{u*} = \begin{cases} \min\left(\dfrac{x}{x + e^{(x-0.5)} - 2e^{(x-0.25)} + 1}, \bar{q}\right) & \text{if} \quad 0 \leq x < 0.25 \\ \min\left(\dfrac{4 - 8x}{-8x + 8e^{(x-0.5)} - 4}, \bar{q}\right) & \text{if} \quad 0.25 \leq x < 0.5 \end{cases} \tag{C.15}$$

Now that we have obtained the optimal pollution levels for each of our three hypothetical population distribution, we can use these values to solve for the pollution at location L on the river.

Before the county split:

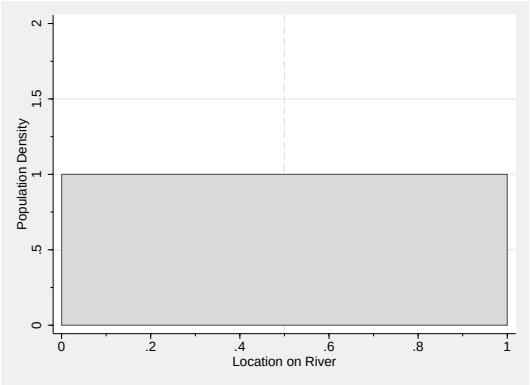
$$P(L) = \int_0^L q_x^* e^{-(L-x)} \cdot f(x) \cdot dx \tag{C.16}$$

After the county split:

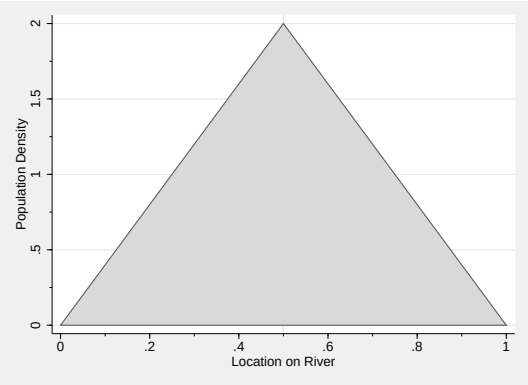
$$P_s(L) = \begin{cases} \displaystyle\int_0^L q_x^{u*} e^{-(L-x)} \cdot f(x) \cdot dx & \text{if } 0 \leq L \leq 0.5 \\ \displaystyle\int_0^{0.5} q_x^{u*} e^{-(L-x)} \cdot f(x) \cdot dx + \int_{0.5}^L q_x^{d*} e^{-(L-x)} \cdot f(x) \cdot dx & \text{if } 0.5 < L \leq 1 \end{cases} \tag{C.17}$$

Since there is no closed form solution for these pollution functions, we numerically integrate these functions in Mathematica and plot them on [0,1]. The graph for the uniform population density function scenario is found in Figure 2 of the paper and the graphs for the scenarios associated with the triangular and triangular-bimodal population densities are displayed in Figures C2 and C3.

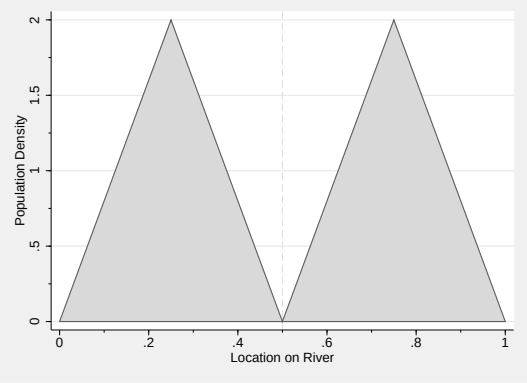
Figure C1: Three Population Density Functions



(a) Uniform



(b) Triangular



(c) Triangular Bimodal

Figure C2: Emissions and Pollution Functions
Triangular Population Density

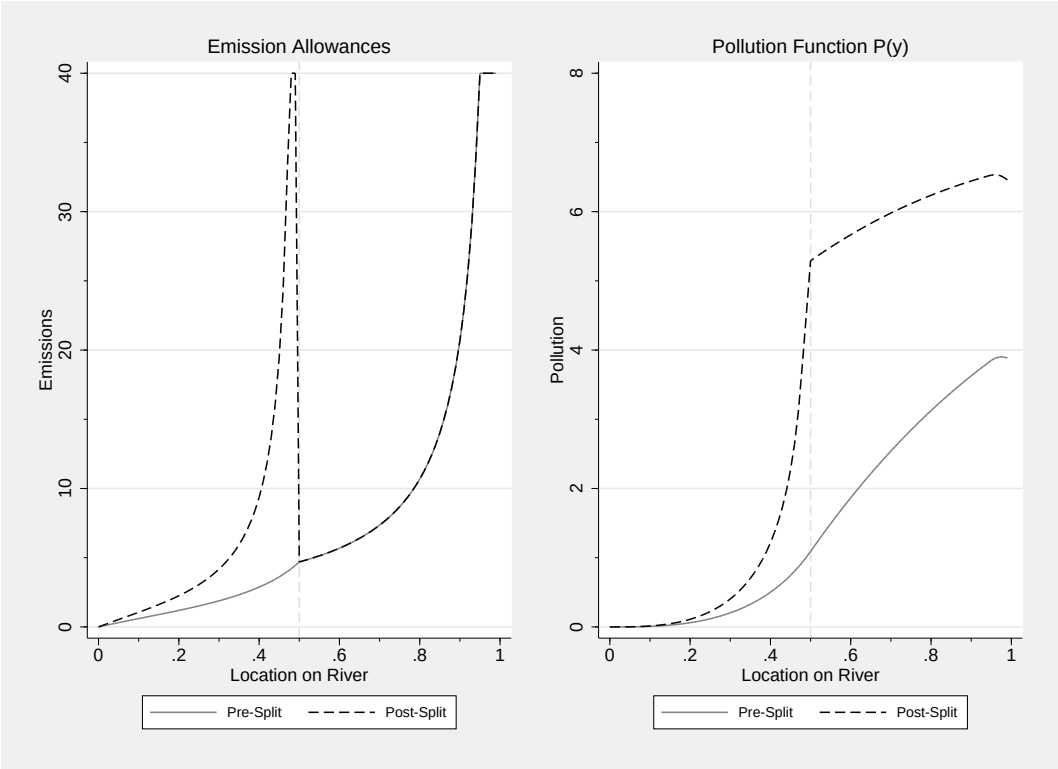
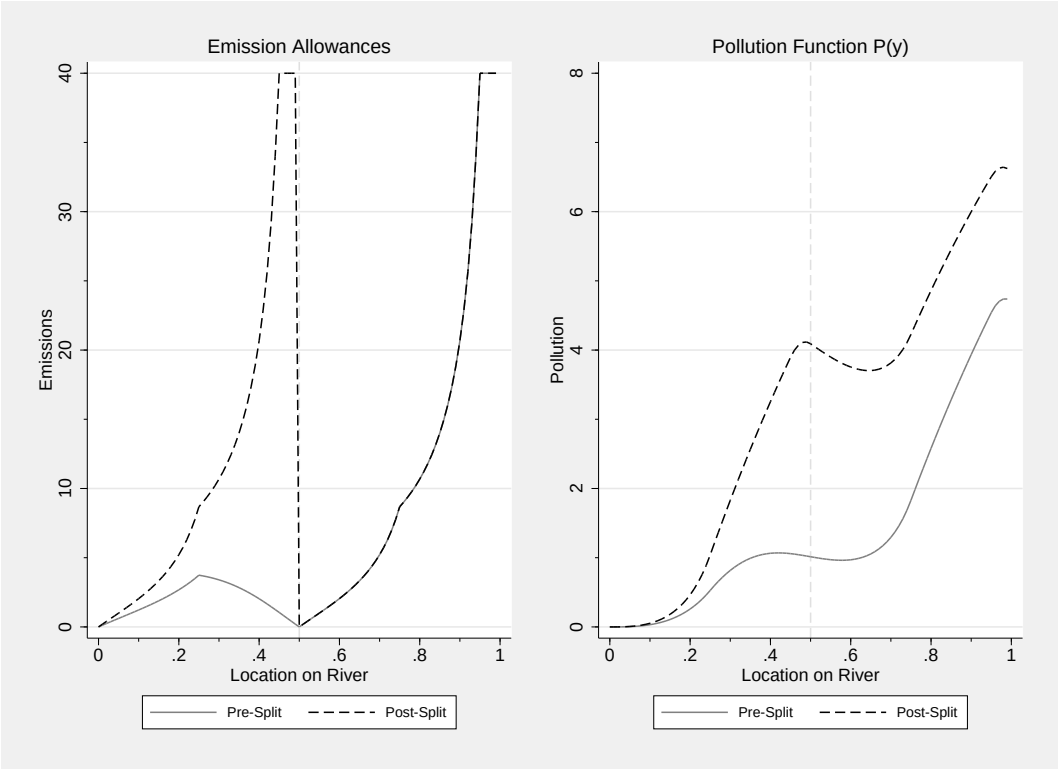


Figure C3: Emissions Allowance and Pollution Functions
Triangular Bimodal Population Density



C.2 Non-Atomistic (Firm or Factory Point Source) Polluters

In this section we model the pollution and emissions functions while restricting emissions to $N = 20$ discrete polluters. In this case, the benefits from emissions are concentrated to these point source polluters. We assume a uniform population density and use the following welfare functions to determine the emissions for a point source polluter at a given location x_n, so that:

$$W_n = ln(q_{x_n}) - \int_{x_n}^{1} q_{x_n} \cdot e^{-(t-x_n)} \cdot d(t) \qquad s.t.\ q_{x_n} < \bar{q} \tag{C.18}$$

$$W_n^u = ln(q_{x_n}^u) - \int_{x_n}^{0.5} q_{x_n}^u \cdot e^{-(t-x_n)} \cdot d(t) \qquad s.t.\ q_{x_n}^u < \bar{q}, \quad 0 \leq x_n \leq 0.5 \tag{C.19}$$

$$W_n^d = ln(q_{x_n}^d) - \int_{x_n}^{1} q_{x_n}^d \cdot e^{-(t-x_n)} \cdot d(t) \qquad s.t.\ q_{x_n}^d < \bar{q}, \quad 0.5 \leq x_n \leq 1 \tag{C.20}$$

where everything is defined as in section B.1.

These welfare functions give rise to the first order conditions:

FOC for (C.18):

$$\frac{1}{q_{x_n}} = \int_{x_n}^{1} e^{-(t-x_n)} \cdot d(t) + \lambda$$

FOC for (C.19):

$$\frac{1}{q_{x_n}^u} = \int_{x_n}^{0.5} e^{-(t-x_n)} \cdot d(t) + \lambda$$

FOC for (C.20):

$$\frac{1}{q_{x_n}^d} = \int_{x_n}^{1} e^{-(t-x_n)} \cdot d(t) + \lambda$$

Solving these equations we derive:

$$q_{x_n}^* = \min\left(\frac{1}{1 - e^{(x-1)}}, \bar{q}\right)$$
$$q_{x_n}^{u*} = \min\left(\frac{1}{1 - e^{(x-0.5)}}, \bar{q}\right).$$

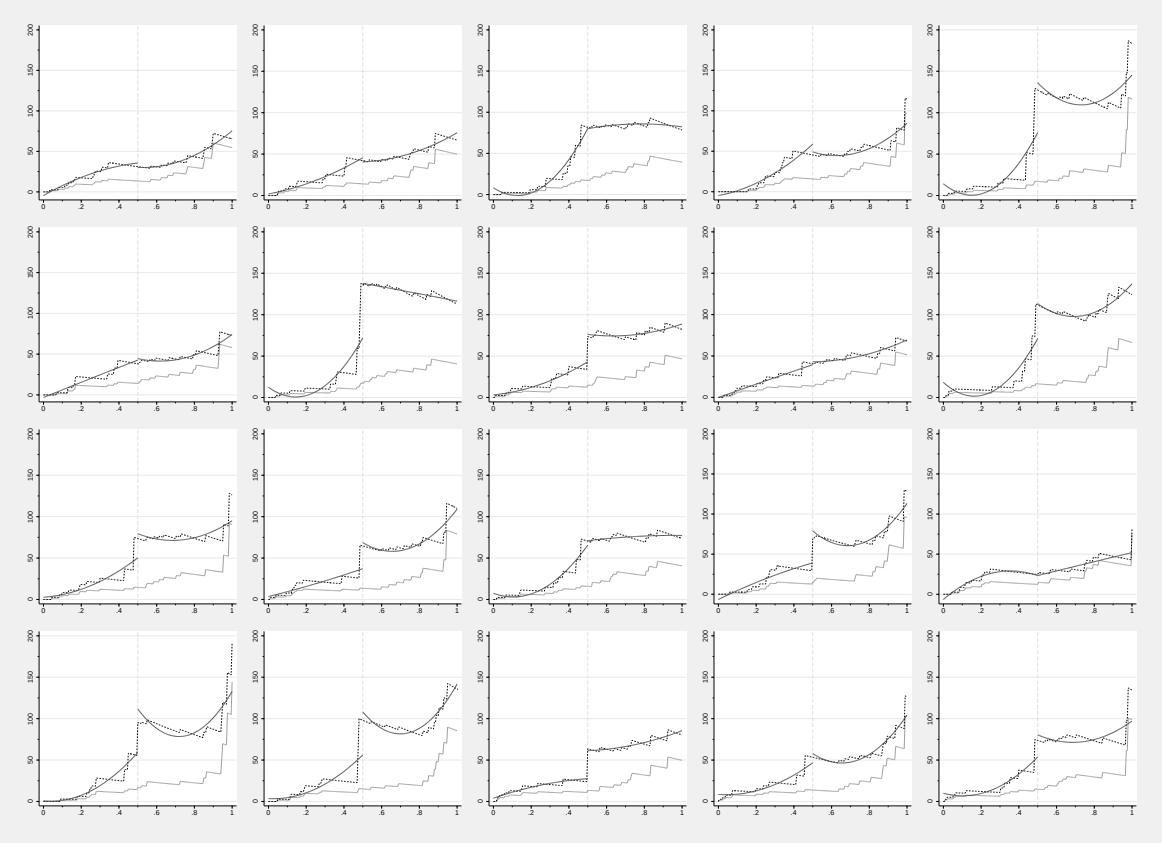
Now that we have obtained the optimal firm-level emissions for any $x_n \in [0, 1]$, we randomly place 10 firms on $[0, 0.5]$ and 10 firms on $[0.5, 1]$. We then compute the pollution functions for the pre-split and post-split situations respectively:

$$P(L) = \sum_{n=1}^{20} q_{x_n}^* e^{-(L-x_n)}$$

$$P_s(L) = \begin{cases} \sum_{n=1}^{10} q_{x_n}^{u*} e^{-(L-x_n)} & \text{if } 0 \leq L \leq 0.5 \\ \sum_{n=1}^{10} q_{x_n}^{u*} e^{-(L-x_n)} + \sum_{n=11}^{20} q_{x_n}^{d*} e^{-(L-x_n)} & \text{if } 0.5 < L \leq 1 \end{cases}$$

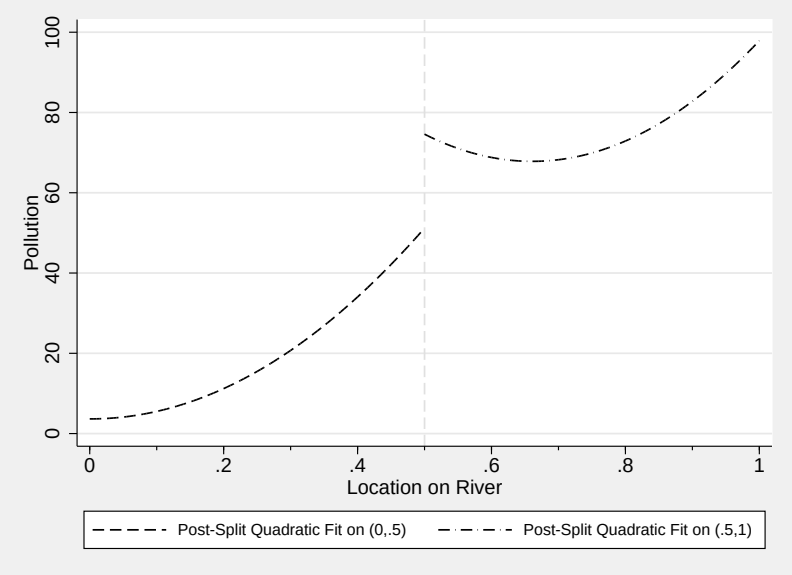
We display the pollution functions for each of our 20 simulations in Figure C4. In each of these cases, we compute a piecewise quadratic fit on [0,0.5] and [0.5,1]. We display the average quadratic fit for all 20 simulations in Figure C5.

Figure C4: Firm Level Pollution Simulations Plot of the Pollution Function $P(Y)$



Note: Y-axis = pollution level; X-axis = river position; population density = uniform. Each simulation conducted by placing 10 firms in $[0, 0.5]$ and 10 firms in $[0.5, 1]$. The light gray lines indicate the pollution level prior to the split, whereas the black dashed lines indicate the pollution level after the split. The solid dark gray lines represent the post-split quadratic fit on $[0, 0.5]$ and $[0.5, 1]$ respectively.

Figure C5: Average Simulated Pollution Function $P(Y)$
Uniform Population Density
Discrete Firms $(N = 20)$



Note: Plot of the function $P(Y) = \beta_0 + \beta_1 DIST + \beta_2 DIST^2$, where the β coefficients are the average of the constant, linear and quadratic terms for the quadratic fit for each of the 20 firm level simulations in Figure C4.

C.3 Partial Internalization of Pollution Costs Imposed on Downstream County

After the county split, it may be the case that the upstream county does not fully discount the effect of their pollution on the downstream county. We derive this version of the model by incorporating a ρ-parameter into our welfare functions that measures how heavily the upstream county weighs the costs they impose on the downstream county. Thus, we obtain the following welfare functions under a uniform population density:

$$W = ln(q_x) - \int_x^1 q_x \cdot e^{-(t-x)} \cdot d(t) \qquad s.t.\ q_x < \bar{q} \tag{C.21}$$

$$W^u = ln(q_x^u) - \int_x^{0.5} q_x^u \cdot e^{-(t-x)} \cdot d(t) - \rho \int_{0.5}^1 q_x^u \cdot e^{-(t-x)} \cdot d(t) \tag{C.22}$$
$$s.t.\ q_x^u < \bar{q}, \quad 0 \leq x \leq 0.5, \quad 0 < \rho < 1$$

$$W^d = ln(q_x^d) - \int_x^1 q_x^d \cdot e^{-(t-x)} \cdot d(t) \qquad s.t.\ q_x^d < \bar{q}, \quad 0.5 \leq x \leq 1 \tag{C.23}$$

These welfare functions give rise to the first order conditions:

FOC for (C.21):

$$\frac{1}{q_x} = \int_x^1 e^{-(t-x)} \cdot d(t) + \lambda$$

FOC for (C.22):

$$\frac{1}{q_x^u} = \int_x^{0.5} e^{-(t-x)} \cdot d(t) + \rho \int_{0.5}^1 e^{-(t-x)} \cdot d(t) + \lambda$$

FOC for (C.23):

$$\frac{1}{q_x^d} = \int_x^1 e^{-(t-x)} \cdot d(t) + \lambda$$

Solving these first order conditions we derive:

$$q_x^* = \min\left(\frac{1}{1 - e^{(x-1)}}, \bar{q}\right)$$
$$q_x^{u*} = \min\left(\frac{e}{(\sqrt{e} - 1)\,\rho \cdot e^x + e - e^{(x+0.5)}}, \bar{q}\right).$$

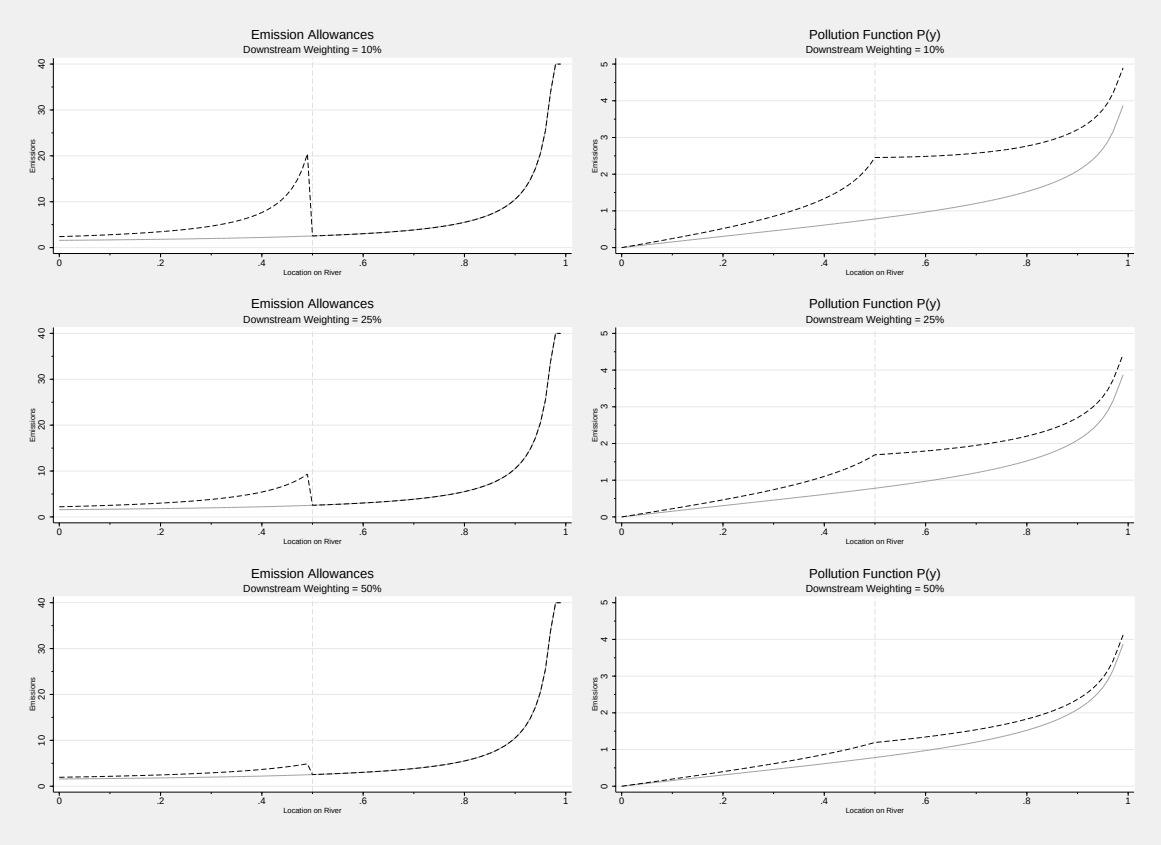
Which gives us the pollution functions:

$$P(L) = \int_0^L q_x^* e^{-(L-x)} \cdot dx \tag{C.24}$$

$$P_s(L) = \begin{cases} \int_0^L q_x^{u*} e^{-(L-x)} \cdot dx & \text{if } 0 \leq L \leq 0.5 \\ \int_0^{0.5} q_x^{u*} e^{-(L-x)} \cdot dx + \int_{0.5}^L q_x^{d*} e^{-(L-x)} \cdot dx & \text{if } 0.5 < L \leq 1 \end{cases} \tag{C.25}$$

We display the emissions and pollution functions for $\rho \in \{0.1, 0.25, 0.5\}$ in Figure C6.

Figure C6: Emissions Allowance and Pollution Functions Assuming Uniform Population Density and Downstream Weighting



Note: The solid grey lines represent the pre-split emissions and pollution levels and the dashed black lines indicate the post-split emissions and pollution levels.

D Implications of Endogenous County Splitting in Densely Populated Areas for Empirical Analysis and Interpretation

As discussed in Section 3 of the paper, a key concern with our estimation strategy to identify the effect of decentralization on pollution spillovers is the possibility of endogenous splitting of counties in areas with high population density (where pollution problems are worsening for an independent reason). Under the assumed form of welfare maximizing behavior by the county authority in our theory, there is actually no such endogeneity problem since the authority would respond to a doubling of the population by simply halving each person's consumption (and pollution) allowance. With twice the population, each person's emissions cause double the harm, so the county authority would force its citizens to cut back on consumption. However, to guide a careful empirical strategy we want to allow for such endogeneity, so we will now assume that each person at location x emits additional unmonitored emissions, ε_x, in addition to their monitored emissions, q_x. The un-monitored pollution increases with population, and is a parsimonious way for us to model endogeneity associated with strategic border placement in high density areas.

We introduce this ε-type pollution into our model by simply modifying the pollution function found in Equation (C.17) to include a ε-term so that our pollution function after the county split is now defined as:

$$P_s(L) = \begin{cases} \displaystyle\int_0^L (q_x^{u*} + \varepsilon) \cdot e^{-(L-x)} \cdot f(x) \cdot dx & \text{if } 0 \leq L \leq 0.5 \\ \displaystyle\int_0^{0.5} (q_x^{u*} + \varepsilon) \cdot e^{-(L-x)} \cdot f(x) \cdot dx + \int_{0.5}^L (q_x^{d*} + \varepsilon) \cdot e^{-(L-x)} \cdot f(x) \cdot dx & \text{if } 0.5 < L \leq 1 \end{cases} \tag{D.1}$$

D.1 Implications of Endogenous County Splitting in High Density Areas

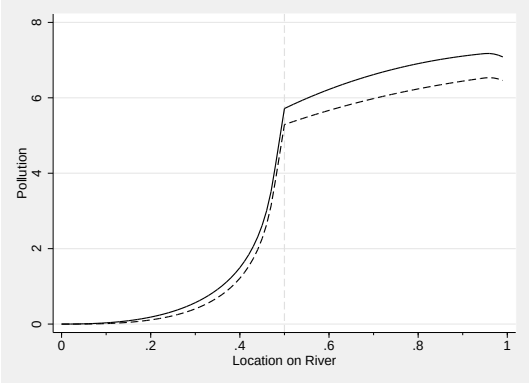
We first model the endogeneity associated with the assumption that counties are more likely to split at locations with higher population density. We relax the assumption of uniform population density and now assume that the population density takes the form of the symmetric triangular distribution found in Equation (C.10). We will examine the effect of a county split at the location coincident with the peak of the distribution (at 0.5), since this is the form of endogeneity of greatest concern (i.e. that splits occur in areas where ε-type pollution increases for independent reasons).

The dashed curves in Figure D1 plot equation (D.1) with $\varepsilon = 0$ and q_x^* and q_x^{u*} (i.e. the optimal per-person regulated emissions) as defined in equations (C.11) and (C.12). The dashed lines thus represent the function which only accounts for 'strategic' q-type pollution. The solid black curves in Figure D1 incorporate the endogenous ε-type pollution using six levels of ε. These lines represent the aggregate of the monitored q-type pollution and the unmonitored ε-type pollution. These curves corresponds to the observed pollution levels in our data, since the data we observe is "total pollution" aggregated across q-type and ε-type.

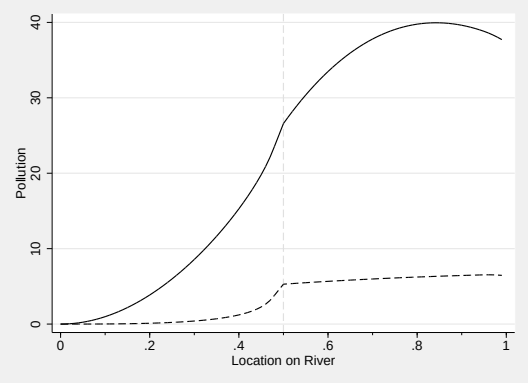
Since ε is the independent effect of population density on pollution that has nothing to do with strategic behavior, increasing values of ε correspond to assuming larger amounts of "endo-

geneity" in our empirical analysis (i.e. that our regressions merely pick up fluctuations in pollution caused by population density changes that have nothing to do with strategic spillovers). We find that the pollution function increases sharply on [0,0.5] for all values of ε because of (i) the strategic motive to pollute more and (ii) the effects of increasing population density which coincide with locations nearby the split. However, once we cross the border and enter the downstream county, we find that the function plotting total pollution ($q+\varepsilon$) continues to increase for any value of ε. This behavior in the downstream county is distinct from the predictions of our baseline model of strategic pollution. This occurs because under the ε model, the border area of the downstream county is still a high-density area, and there are more polluters in this region. This observation gives rise to an empirical test of the quantitative importance of this type of endogeneity, since it implies that the coefficient on $U1$ in equation (4) (i.e. the slope of the pollution function downstream of a border) should be positive. However, in our empirics the coefficient estimates for this variable are consistently negative, which indicates that the empirical results reported in this paper are likely evidence of strategic behavior as opposed to spurious correlation due to borders being drawn in areas of high density. As the value of ε increases in D1, the slope $U1$ becomes more and more positive, and the discontinuity in slope at the border disappears. This suggests that Prediction 3 from the baseline theoretical model (the discontinuity in the slope of the pollution function at the border) provides a clear empirical test of the possibility of endogenous county-splitting in high density areas. If our main pollution results are primarily driven by endogenous splitting in high-density, high-pollution areas, then we would not observe the structural break in the slope of the pollution function at the border.

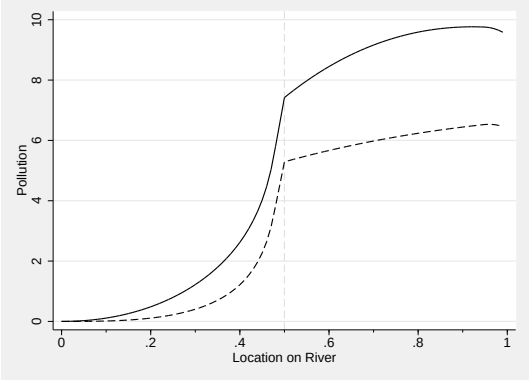
Figure D1: Endogenous Population Density Based Split under a Triangular Population Distribution: Effect of Varying Levels of Unmonitored ε-type Pollution



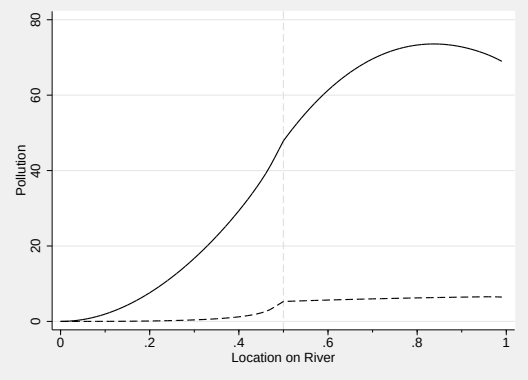
(a) $\varepsilon = 1$



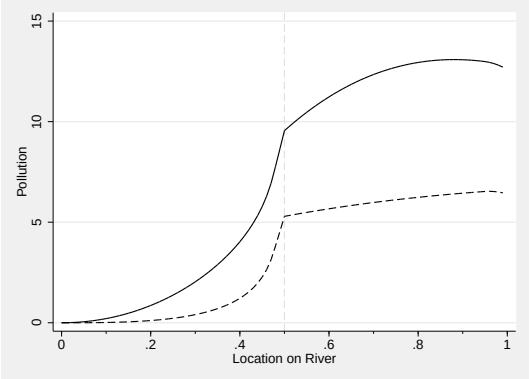
(b) $\varepsilon = 50$



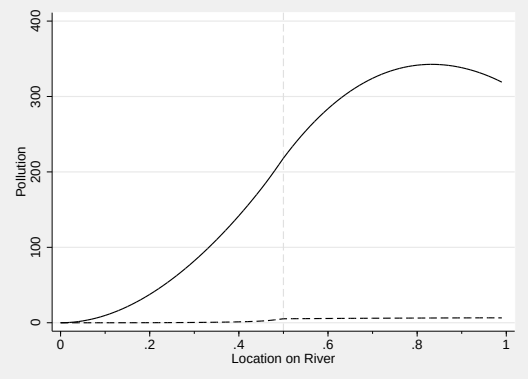
(c) $\varepsilon = 5$



(d) $\varepsilon = 100$



(e) $\varepsilon = 10$



(f) $\varepsilon = 500$

Note: The dotted lines indicate the pollution function prior to including ε-type pollution. The solid lines represent the pollution function after the inclusion of ε-type pollution. The triangular population density function is defined in Equation (C.10) and is plotted in Figure C2c.

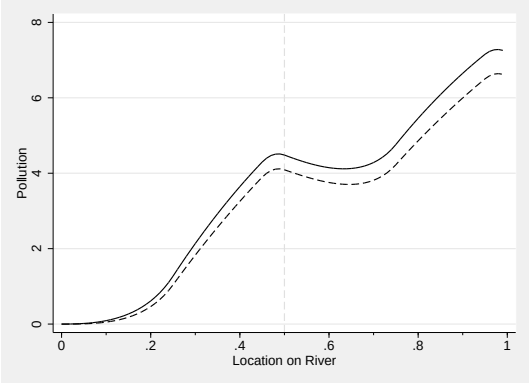
D.2 Implications of Endogenous County Splitting in Low Density Areas

We also want to examine the empirical implications of another plausible form of endogeneity – that two distinct areas of the original county experience population growth, and that the new border is drawn in the outlying low density area that falls in between those two areas of population growth. We model this possibility using the bimodal symmetric triangular distribution found in equation (C.13), and the resulting pollution functions are shown in Figure D2. We plot the pollution function, Equation (C.17), for various levels of ε with the optimal per-person emissions for the triangular bimodal distribution found in Equations (C.14) and (C.15). As Figure D2 shows, an important and distinctive implication of this form of endogeneity is that the rate of increase in pollution slows down as we approach the newly drawn border. In fact, an inflection point in this function for any value of ε occurs at $x = 0.25$, where the pollution function begins to increase at a decreasing rate, since population density decreases as we approach the border. This implication is in direct contrast with prediction 2 from the theory - that pollution should increase at an increasing rate as we approach a downstream exit border.

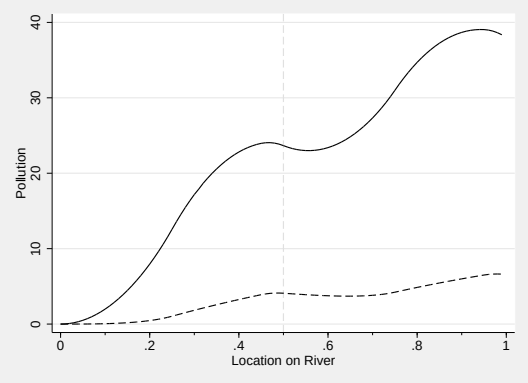
We add a non-linear (quadratic) term for distance to the downstream border in our estimating equation ($1D^2$ in equation 4), and the coefficient on this term provides an empirical test of this form of endogeneity. If borders are being drawn in low-density areas, then we should observe a negative coefficient on $1D^2$. We end up estimating a positive and significant coefficient on $1D^2$, which suggests that the data are not consistent with this specific form of endogeneity. We have also experimented with a cubic specification (results available on request), but even with this greater flexibility, the empirically estimated pollution function does not display a decreasing slope upstream of the exit border.

In summary, the set of four congruent predictions and empirical results that we report in this paper are not easily explained away by either of two specific plausible alternative hypotheses.

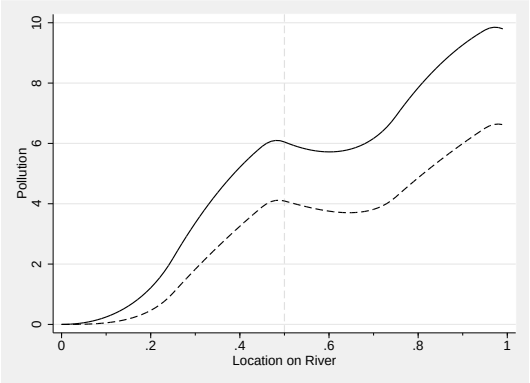
Figure D2: Endogenous Population Density Based Split under a Triangular Bimodal Population Distribution: Effect of Varying Levels of Unmonitored ε-type Pollution



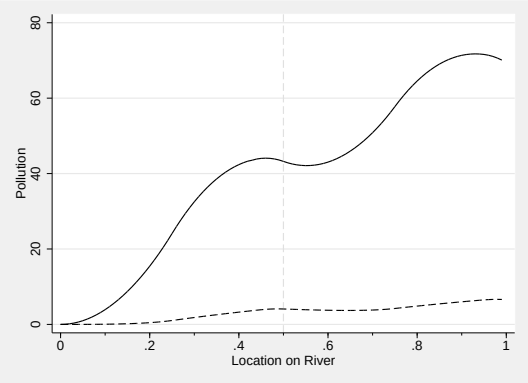
(a) $\varepsilon = 1$



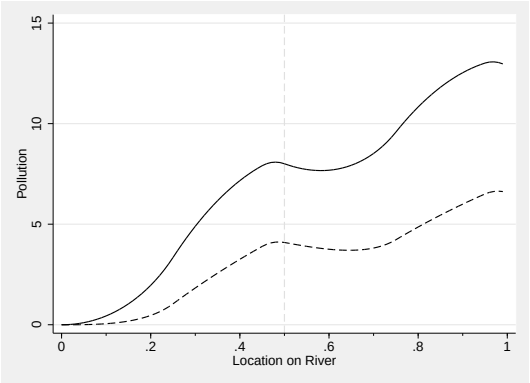
(b) $\varepsilon = 50$



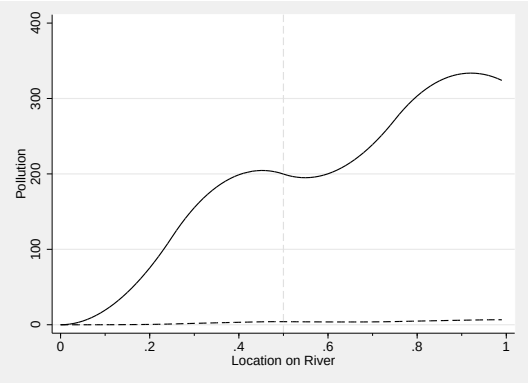
(c) $\varepsilon = 5$



(d) $\varepsilon = 100$



(e) $\varepsilon = 10$



(f) $\varepsilon = 500$

Note: The dashed lines indicate the pollution function prior to including ε-type pollution. The solid lines represent the pollution function after the inclusion of ε-type pollution. The triangular bimodal population density function is defined in Equation (C.13) and is plotted in Figure C2b.

E Chemical Properties of Biochemical Oxygen Demand

Biochemical Oxygen Demand (BOD) measures the amount of decomposable organic matter in the water. During the decomposition process, micro-organisms feed on organic matter until it is processed into an inorganic form. BOD measures the amount of organic matter in the water indirectly through measuring the amount of oxygen demand for a water sample over 5 days (Tchobanoglous and Schoeder, 1985).

BOD increases as micro-organisms accumulate to degrade organic material. High levels of organic pollution and BOD are associated with river eutrophication (Chapman et al., 1996, p. 276–278).

Organic pollution may be derived from a variety of sources. Common organic pollutants include: phenols, which are common in industrial food manufacturing; surfactants, which are a by-product of detergents and are common in both household and industrial waste; sewage; agricultural and urban run-off; and domestic waste. Industries which emit pollutants to which BOD levels are particularly sensitive include: food processing, oil extraction and refining (sugar cane refining is a particularly large industry in Brazil, and untreated waste waters from sugar refineries carry high organic pollution loads), pulp and paper industries, and textiles. BOD is also sensitive to pollutants from chemical and pharmaceutical industries, mining, metallurgy, and machine production (Chapman et al., 1996, p. 122).

Water Type	BOD Level
Unpolluted	2
Highly Polluted	10
Treated Sewage	20–100
Raw Sewage	600
Industrial Waste	up to 25,000 †

BOD is an approximation of theoretical oxygen demand, or the total oxygen which would be necessary to decompose the organic matter present in the sample. It is measured as the oxygen consumption in a given water sample at twenty degrees Celsius over a period of five days. Consumption is determined as the difference in dissolved oxygen content between the beginning of the incubation period and at the end of five days. A gestation period of five days is given as the oxygen consumption of the micro-organisms is initially high, but decreases as organic pollutant concentrations decrease (Hounslow, 1995, p. 302).

Attenuation rates of organic pollution depend on a host of local factors. Weather can affect decomposition rates as low temperatures increase the half-life of organic pollution and slow the process of decomposition. High levels of water evaporation may increase the concentration of organic pollution while increased rainfall may contribute to the dilution of the pollution loads. High levels of rainfall may, however, also lead to local flooding and increased contamination from erosion. Geological factors such as local soil and rock types affect the absorption of pollutants into the river bed. Geographical factors such as slope, elevation, discharge, and depth affect attenuation since water velocity increases churning and oxygenation leading to faster oxidation periods for organic pollutants (Chapman et al., 1996, p. 246–276).

† Chapman et al. 1996, p. 88

F Analysis of Light Density Data

A network data set of stream centerlines was created from the Hydro 1K USGS digital elevation dataset. Upstream and downstream terminal nodes corresponding to each stream's head and mouth were extracted and stored as point features. The default precision of the network geoprocessing tools was then adjusted to avoid an address overflow error in ArcGIS that caused NULL values to be stored for calculated linear referencing distances. Routes were created from each headwater node to the corresponding mouth node, and selected such that only the longest stream (for each monitoring station) intersecting at least two monitoring stations was considered.

Defense Meteorological Satellite Program-Operational Line Scan (DMSP-OLS) Nighttime lights rasters for the years 1992 to 2012 were masked to an area within a 1 km buffer of the selected streams. The masked raster geometry was converted to individual point features corresponding to the locations of the pixel centers. These points were used to extract digital number values from the entire set of DMSP-OLS nighttime lights rasters (all years and all satellites).

In ARCGIS, routes along the rivers were dissolved to a single polyline feature and used to create a river network of the longest streams intersecting two or more monitoring stations. Points corresponding to raster cell centers were transformed from the native geographic reference system of the nighttime lights data (WGS84) to the projection of the streams data (South American Datum 1969 Equidistant Conic), and entered as "Incident" locations in an ESRI ArcGIS Network Analyst Closest Facility analysis. Downstream terminal nodes for each stream served as "Facility" locations in this analysis. Facility searches were restricted to the downstream direction. Network distance from each raster cell location to the stream mouth was stored in the "Incident" points. An additional OD Cost Matrix analysis was performed in order to determine the nearest downstream monitoring location and its distance to each raster cell. The direction of the network was then reversed, and the nearest upstream monitoring station (and distance) was found for each raster cell.

Due to the high data volume (456,739 raster cells after masking), the network analyses described above were broken into batches of 5000 points, with each result stored in a separate table. A Python script controlled the creation of each batch and the execution of each set of analysis. The result tables were merged and joined to the raster cell points in the original WGS84 projection. The upstream distance values found for each raster cell point were used to create a raster of distances from the mouth of each stream. Zonal statistics were found for the raster cells based on municipio boundaries obtained from the Instituto Brasileiro de Geografia e Estatística (IBGE) such that the minimum and maximum upstream distance in each municipio in each year were stored in each cell location.

The relative distance upstream for each raster cell in each municipio was determined using the raster calculator by $(D - Dmin)/(Dmax - Dmin)$. Relative upstream distances for each of the municipio boundary datasets were extracted to the point features created from the raster cell centers, and point features were spatially joined with the municipio boundaries from each redistricting. Points with municipio codes that changed between successive years were flagged as split municipios when new municipio codes were found in different years of the municipio boundaries dataset.

G Robustness Checks of Water Pollution Results

A few pollution monitoring stations are located close to areas where sewage is emitted directly into the water, and BOD is measured to be extremely high at these locations. To check whether the pollution patterns we report are driven by these outliers, we re-run the station-level and station-pair-level regressions either omitting the top 1%, 3% or 5% of BOD values, or top-coding those extreme observations at the 99th, 97th or 95th percentile of BOD values in the sample. Table G1 shows these results for the station-level regressions (analogous to Table 2). The evidence in favor of predictions 1, 2 and 3 remain equally strong. Pollution increases by 2-3% per kilometer as river approaches the downstream exit border, and the coefficient on 1D2 indicates that the rate of pollution increase is faster closer to the border. The slope of the pollution function switches to -3.6% to -5.1% at the border, and this change in slope remains statistically significant.

Table G2 shows that the evidence in favor of prediction 4 is not sensitive to the exclusion of outliers either. Each additional border crossing is associated with a 3-4% increase in water pollution. In this station-pair specification, pollution increases by 1.5% every kilometer closer the river gets to the downstream exit border, and the slope of the pollution function switches to -3% on the other side of the border.

Table G1: Robustness Check: Station-Level Regression

	Dependent Variable: 100*LogBOD in the upstream station			
	Top 1% dropped		Top 1% top coded	
Distance from station 1 to Downstream border (1D)	-2.323***	-3.297***	-2.056***	-3.130**
	(0.848)	(1.247)	(0.776)	(1.254)
Squared Distance from station 1 to Downstream border ($1D^2$)	0.035***	0.039**	0.033***	0.037*
	(0.012)	(0.020)	(0.011)	(0.019)
Distance from upstream border to station 1 (U1)	-4.276	-4.926	-3.659	-5.059
	(2.997)	(5.181)	(3.073)	(5.241)
Squared Distance from upstream border to station 1 ($U1^2$)	0.238	0.411	0.204	0.411
	(0.185)	(0.372)	(0.188)	(0.375)
GDP of the county in which the upstream station is located in 100000 Reis	0.677***	0.220	0.409	0.002
	(0.184)	(0.192)	(0.318)	(0.551)
Population of the upstream county, in 100,000	-3.638	1.263	0.312	3.197
	(2.387)	(4.592)	(2.048)	(6.020)
Area of the upstream county, in 100,000 square km	-206.116*	-226.390*	-252.729**	-251.029**
	(107.581)	(128.719)	(110.863)	(122.854)
Includes station pair trends	N	Y	N	Y
Observations	5,940	5,940	5,989	5,989
R-squared	0.070	0.137	0.067	0.135
Number of pairs	372	372	372	372

The dependent variable is 100*the log level BOD at the upstream station. All regressions include station pair fixed effects. Standard errors are clustered by the station. All regressions include an indicator variable for stations which are not separated by a border, and the distance 1D and U2 (the distance from station 1 to the nearest downstream border and the distance from the nearest upstream border to station 2 are set equal to 0). An indicator variable is also included for cases in which there were no intermediate counties for GDP variables (cases of 0 or 1 crossings).

Table G2: Robustness Check: Station-Pair Regression

	Dependent Variable: 100* (Log downstream BOD-Log Upstream BOD)			
	Top coded 1% of observations		Top 1% of observations dropped	
Number of borders crossed between station 1 and station 2	3.522**	3.639**	2.949**	3.063**
	(1.518)	(1.502)	(1.328)	(1.315)
Distance from station 1 to Downstream border (1D)	1.452	1.444*	1.594*	1.591**
	(0.898)	(0.863)	(0.820)	(0.785)
Distance from Upstream border to station 2 (U2)	-3.204*	-3.005	-2.908*	-2.710*
	(1.786)	(1.853)	(1.511)	(1.543)
Outside Station Pair Control Variable: Distance from Upstream border to station 1 (U1)		-1.723		-1.639
		(2.209)		(2.061)
Outside Station Pair Control Variable: Distance from station 2 to its Downstream border (2D)		-0.127		-0.150
		(0.512)		(0.497)
Observations	5,989	5,989	5,913	5,913
R-squared	0.117	0.117	0.114	0.114
Number of pairs	372	372	372	372
F-test for slope of pollution function upstream of border = slope downstream of border	3.792	3.245	4.931	4.267
Prob > F	0.0528	0.0730	0.0274	0.0400

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

The dependent variable is 100*the log difference in BOD between the downstream station and the upstream station. All regressions include station pair fixed effects, station pair trends, river basin-year, and river basin month dummies. All regressions also include controls for GDP, population, and area in the upstream county, the downstream county, and the average in the intermediate counties. Standard errors are clustered by the downstream station. All regressions include an indicator variable for stations which are not separated by a border, and the distance 1D and U2 (the distance from station 1 to the nearest downstream border and the distance from the nearest upstream border to station 2 are set equal to 0. An indicator variable is also included for cases in cases in which there were no intermediate counties for GDP variables (cases of 0 or 1 crossings). In the first three regressions BOD observations for both the upstream and downstream stations above the 1 percentile observation are coded at that observation (41.7).

H Formation of Basin Committees

In recent years, Brazilian Federal and (a number of) state governments have recognized the need for coordinated management of water resources, as decentralized areas experienced highly uneven enforcement of pollution norms. The concept of "Water basin management" was developed with the expectation that it would involve all relevant stake-holders within the basin (who bears any externality costs stemming from water use by one party) and address the need for coordination at that level, and democratize decision making (Abers et al., 2009).

Water basin committees are charged with determining and implementing water use charges, setting expected water quality, creating a forum for negotiation between stakeholders along a river, and overseeing the management of the river by approving plans and resolving conflicts. Bulk water charges (cobrança), set by the committee were expected to finance committee expenses, but many committees had not set these rates even several years after inception (Abers and Keck, 2006). These committees have been more active in elaborating water basin management plans, encouraging negotiation between stake-holders, and forming partnerships between the members of the committees (Abers and Dino, 2005). In response to a survey, 53% of representatives of water basin management committees said that the impetus to form the committee came in large part from a conflict over water use and 50% of representatives said that the committee was formed because the conditions of the river were worsening (Abers and Dino, 2005).

State laws differ significantly in their requirements of water basin management committees from federal law, and even within a state there is considerable variation between water basin management committees in terms of their design. Federal law states that the committees should be composed of 40% major water users, 20% small water users and civil society, and 40% representatives of the state, local, and federal government. However, in practice, less than 20% represent major water users, nearly 40% represent civil society, and over 40% represent state, federal and local governments (Abers et al., 2009). The amount of participation of different stakeholders varies across the basins, but municipal governments and large stakeholders tend to devote the most time to the water basin committees, with 82.1% of municipal representatives and 80.5% of large industry representatives spending at least one day per month on the management of the committee (Gutíerrez, 2008).

We have collected information on the 153 water basin committees operating in Brazil and the years in which they have been active. This allows us to estimate pollution functions (and the effects of border crossings and distances, which are the main subject of this paper) separately for areas that have water basin committees and areas that do not. We use this analysis to explore whether functioning basin committees help to mitigate the cross-border pollution externalities documented in this paper.

Furthermore, we use municipio-level electoral data from IPEA and the Brazilian Election Commission to identify election cycles during which the mayor of the county where the downstream station is located has the same party affiliation as the mayor of her neighbor county immediately upstream. A match in party affiliation is potentially an alternate mechanism by which the transaction costs between municipios falls, and negotiation over water quality becomes more feasible. We therefore also estimate pollution functions as rivers cross county borders, separately for the sub-sample in which the parties in power in the upstream and the downstream counties match in terms of political affiliation, and the sub-sample in which they do not.

References

Abers, R. N. and Dino, K. J. Decentralizçáo da gestao da agua: Por que os comites de bacia estao sendo criados? *Ambiente e Sociedade*, 8(2), 2005.

Abers, R. N., Formiga-Johnsson, R. M., Keck, M. E., Frank, B., and Lemos, M. C. Inclusão, deliberação e controle: três dimensões de democracia nos comitês e consórcios de bacias hidrográficas no brasil. *Ambiente & Sociedade*, 12(1):115–132, 2009. URL `http://www.scielo.br/pdf/asoc/v12n1/v12n1a09.pdf`.

Abers, R. N. and Keck, M. E. Muddy waters: The political construction of deliberative river basin governance in brazil. *International Journal of Urban and Regional Research*, 30(3):601–622, 2006. URL `http://dx.doi.org/10.1111/j.1468-2427.2006.00691.x`.

Affonso, R. Os municípios e os desafios da federação no brasil. *São Paulo em Perspectiva*, 10(3):3–10, 1996.

Bremaeker, F. E. J. Limites à criação de novos municípios : A emenda constitutional no. 15. *Revista de Administração Municipal*, 43(219), 1996a.

Bremaeker, F. E. J. Os novos municípios: Surgimento, problemas e soluções. *Revista de Administração Municipal*, 40(206):88–99, 1996b.

Cachatori, T. L. and Cigolini, A. A. Emancipações municipais no brasil: Prognóstico sobre a continuidade da compartimentação do espaço em novos municípios. *Revista Geonorte, Edição Especial 3*, 7(1):730–747, 2013.

Chapman, D. V., Organization, W. H., Press, C. R. C., and undefined, o. *Water quality assessments: a guide to the use of biota, sediments and water in environmental monitoring*. E & Fn Spon London, 1996. URL `http://wwwlive.who.int/entity/water_sanitation_health/resourcesquality/watqualassess.pdf`.

Gomes, G. M. and MacDowell, M. C. Descentralização política, federalismo fiscal e criação de municípios: O que é mau para o econômico nem sempre é bom para o social. In *Texto para Discussão, Instituto de Pesquisa Econômica Aplicada*. 2000.

Gutíerrez, R. Governo municipal e gestão de bacia hidrografica no brasil. *Seminario Auga da Gente*, 2008.

Hounslow, A. *Water quality data: analysis and interpretation*. CRC press, 1995. URL `http://books.google.com/books?hl=en&lr=&id=scQ9RCHni8kC&oi=fnd&pg=PA1&dq=Hounslow+1995&ots=hOpdykyJXx&sig=sg2WUPC2YitvsGifXEEREP9x7GA`.

Lordello de Mello, D. O municipio na organização nacional; bases para uma reforma no regime municipal brasileiro. *Rio de Janeiro: IBAM*, 1971.

Lordello de Mello, D. A multiplicação dos municípios no brasil. *Revista de Administração Municipal*, 39(203):23–28, 1992.

MacLachlan, C. M. *A History of Modern Brazil: The Past Against the Future*. Rowman & Littlefield, 2003.

National Congress of Brazil. Proposta de lei complementar. Technical Report 416, 2008.

Noronha, R. d. Criação de novos municípios: o processo ameaçado. *Revista de Administração Municipal*, 43(219):110–117, 1996.

Noronha, R. d. and Cardoso, E. D. Emancipações municipais: Como ficam os municípios de origem? *Revista de Administração Municipal*, 42(214):67–80, 1996.

Tchobanoglous, G. and Schoeder, E. D. *Water Quality: Characteristics, Modeling and Modification*. Prentice Hall, Reading, Mass, 1 edition edition, 1985.