Decentralization and the Political Economy of Water Pollution: Evidence from the Re-drawing of County Borders in Brazil*

Molly Lipscomb
University of Notre Dame, Department of Economics

Ahmed Mushfiq Mobarak
Yale University, School of Management

Abstract

Using panel data on water pollution in Brazilian rivers, we study inter-jurisdictional externalities when more jurisdictions start managing the same river segment. Counties split at each election cycle, which change the locations of borders. This allow us to identify the effects of border crossings and the spatial pattern of pollution as rivers approach borders, controlling for fixed effects and trends specific to each location. The theory of externalities predicts that pollution should increase at an increasing rate as the river approaches the downstream exit border, that there should be a structural break in the slope of the pollution function at the border, and that larger number of managing jurisdictions should exacerbate pollution externalities. We find support for all four predictions in the data. We also show that it is difficult to reconcile other theories and mechanisms with the four pieces of congruent evidence we report. These results suggest that decentralization policies promoted by governments and international organizations come at an efficiency cost when local jurisdictions provide public services with inter-jurisdictional spillovers. We provide this evidence in an important setting: over one billion people in the world lack sufficient potable water, and 80 percent of sewage is dumped into surface water untreated.

JEL Codes: Q56, D62, H23, Q53, O12

Keywords: Decentralization, Water Pollution, Inter-jurisdictional Externalities

November 22, 2011

---

* First Draft: October 2011. An older version of this paper with less data circulated in 2007 as "Decentralization and Water Pollution Spillovers." We thank Marianne Bertrand, Erin Mansur, Bernardo Mueller, Jim Berry, Bill Evans, Andrew Foster, Claudio Ferraz, numerous water management practitioners at the federal and various state water agencies in Brazil, and seminar participants at Harvard University, University of Michigan, Yale University, the University of Colorado at Boulder, Resources for the Future, NC State, Wesleyan University, the World Bank Microeconomics of Growth Conference (Rio de Janeiro), 2008 Econometric Society Winter Meetings, 2007 NEUDC Conference, 2007 NBER Summer Institute in Environmental Economics, 2007 BREAD conference and the 2006 ISNIE conference for helpful discussions and comments. We are grateful to *Resources for the Future*, *Corporacion Andina de Fomento* and the *National Science Foundation* for financial support, and to the *World Resources Institute* for map data. All errors are our own.

Correspondence: Mushfiq Mobarak, ahmed.mobarak@yale.edu, +1 203-432-5787

1. Introduction

Many development institutions and practitioners promote decentralization as a way to improve governance and targeting of services to the poor. The idea featured prominently in the World Bank's flagship publication on public service delivery, which notes that "Decentralizing delivery responsibilities for public services is prominent on the reform agenda in many developing countries. Bolivia, India, Indonesia, Nigeria, Pakistan, and South Africa - to name a few - are all part of a worldwide movement to decentralize." (World Bank 2004). The merits of decentralized decision-making have been debated in the public finance and environmental economics literatures (Oates 1972, Cumberland 1981, Oates and Schwab 1988, Levinson 1997, Besley & Coate 2000, Bardhan and Mookherjee 2000, List and Mason 2001), which contend on the one hand that decentralization can improve targeting when there is local variation in preferences, while on the other hand that it can generate externalities across jurisdictional boundaries.

This paper empirically examines the nature and size of externalities from decentralizing decision-making in a setting in which spillovers are important and their effects transparent and easy to measure: pollution in rivers as they cross jurisdictional boundaries. Upstream and downstream interest groups created by the flow of rivers across borders have led to hundreds of international and intra-national conflicts over water use and management (Wolf 2002).[1] With negative spillovers on downstream users, decentralizing water

---

[1] Recent international disputes over water quality and quantity include conflicts over water rights in the Jordan River (which have aggravated the Israel-Palestine conflict (Mustafa, 1994)), disagreements over dams in the Euphrates between Turkey and Syria (Jongerden, 2010), and contention over irrigation canals between Pakistan and India (Wolf, 1998). Intra-national conflicts over use of the Colorado River by Upper- and Lower-basin states have divided the Western U.S. for over 100 years, and have extended to include rural-urban conflicts over water rights as well as conflict between state and federal governments and Native American populations (Postel, 1999; Hobbs, 2005). Similarly, disputes have arisen between Indian states over diversion of water between watersheds and preferential hierarchies of water uses (Richards and Singh, 2002).

management may be inefficient from a societal perspective in the absence of inter-jurisdictional coordination.

To examine externalities across jurisdictions we compile a rich panel dataset of 5989 water quality measures collected at monthly intervals at 372 upstream-downstream pairs of monitoring stations located in all eight major river basins across Brazil. Our unique empirical strategy overlays these pollution data on a series of GIS maps of county (*municipio*) boundaries which shift and evolve over time due to county-splitting. Brazil redraws county borders frequently (the number of counties increased from 3991 in 1980 to 4491 in 1990 to 5507 in 2000), which changes both the number of border crossings between a pair of water quality monitoring stations and the distance of each station to its nearest upstream and downstream borders over time. This is turn changes the incentive to pollute at a given location, since some downstream constituents become "external" to the upstream politician's calculus after the re-districting.

The change-in-border based identification strategy we use is critical, because identifying externalities based on cross-sectional comparisons across different rivers would be subject to a number of bias concerns. Geographic and demographic differences between small and large jurisdictions (with more or less frequent border crossings), endogenous placement of monitoring stations, and political differences between counties could all generate omitted variable biases in cross sectional estimates of the effect of pollution near borders. We present evidence of strategic pollution by jurisdictions using only inter-temporal *changes* in water quality at a specific location as borders move closer or as the number of border crossings change, while controlling for location fixed effects.

We model the behavior of a local policy-maker in a decentralized decision-making environment who cares about his own constituency in optimally allocating production (and pollution) along a river that travels through his jurisdiction, but does not account for the external effects on downstream jurisdictions. The theory makes four specific predictions that we test with our data: (1) Pollution increases as the river travels towards the downstream border (since there is less harm caused on own constituents by polluting farther downstream), (2) the pollution increases at an increasing rate as the river travels downstream, (3) there is an inflection or structural break in the slope of the pollution function at the border, as the new jurisdiction does not allow new emissions at the same high rate at its own most upstream locations, and the pollution traveling into the county attenuates with the flow of water, and (4) pollution along a river increases with a larger number of border crossings.

We estimate the entire shape of the pollution function implied by this theory, and find support for all four predictions. For every kilometer closer a river is to a downstream border, Biochemical Oxygen Demand (the pollution measure) increases by 1.5-3%. This increase occurs at a faster rate as the river gets closer to the downstream border, and the pollution function therefore has the steepest positive slope closest to the river's exit point out of the jurisdiction. We also identify a structural break in the slope of the pollution function at the border, which suggests that the county that the river enters restricts polluting activity in its own upstream areas. Finally, for each additional border crossing induced by a county redistricting, pollution increases by over 3%.

The last finding builds on the results of Sigman (2002, 2005)'s seminal analyses of pollution in international rivers and in rivers in the United States. Sigman (2002) shows that water pollution is larger when a river crosses an international border. However, as that paper notes, proximity to borders may be correlated with unobserved heterogeneity in geography, population density or economic activities. We demonstrate the existence of such pollution spillovers within a country using a more stringent design controlling for location fixed effects.

Sigman (2005) shows that decentralization of environmental authority to the state level in the United States led to a 4% degradation of water quality downstream of authorized states, with an environmental cost downstream of $17 million annually. This analysis includes monitoring station fixed effects to account for unobserved heterogeneity, but since borders are fixed, it is not possible to study the effects of proximity to upstream and downstream borders in that setup. In contrast, our approach allows for analysis of the pollution function at varying distances from the border, and we estimate the non-linear shape of the pollution function on both sides of the border. This has an important econometric advantage since placement of water quality monitoring stations may not be random. Monitoring stations are often placed by environmental authorities in areas where water pollution is of particular health or legal concern. In our analysis, we identify the pollution impact of changes in distance to the border *controlling for the fixed location* of the monitoring station, using only the variation resulting from the re-drawing of county borders in Brazil.

A key concern with our border changes-based identification strategy is that factors correlated with pollution, such as increases in population density, can affect a county's propensity to split. The fact that we estimate the non-linear shape of the entire pollution function on both sides of the border is helpful in this regard, since time-varying unobservables cannot easily explain the specific pattern of strategic pollution shifting we report. Nonetheless we want to be as careful as possible in differentiating evidence of true strategic behavior from spurious correlations, and therefore report results controlling for monitoring station pair specific linear trends. Increasing population density or economic activity are gradual relative to the sharp county splitting "events" every four years associated with the election cycle. The station pair trends account for gradual changes, and these specifications identify the pollution spillover effects using only variation that corresponds to these sharply dated events.

We also theoretically model specific forms of endogeneity (where a jurisdictional split occurs in an area with high population density, for example), and examine whether or not the spatial pattern of pollution implied by the theory is consistent with our estimated pollution function. These alternative theories are difficult to reconcile with the structural break that we observe in the slope of the pollution function at the border.

Brazilian counties' strategic pollution close to the river's downstream point of exit from a county serves as a cautionary tale for decentralization initiatives without adequate inter-jurisdictional coordination mechanisms in place.[2] Our results thus contribute to the academic and policy debate on the effects of political decentralization.[3] We provide empirical evidence on the inter-jurisdictional spillover aspect of this issue using a novel research design, and contribute to a growing empirical literature (Faguet 2004, Millimet 2003, Foster and Rosenzweig 2002, List and Mason 2001) that has struggled with the econometric difficulties posed by the endogenous selection of decentralization initiatives.

---

[2] In fact, Brazil has begun to recognize the problem we document, and has encouraged the formation of river basin committees to improve negotiation between the relevant stake holders. This has been encouraged first at the state level, by the state of Sao Paulo in 1991, and later at the federal level through the 1997 National Water Law. As of 2007, more than 100 river basin committees had been created (Abers, 2007).

[3] We examine the effect of a very specific form of decentralization--the geographic splitting of counties leading to a larger number of counties managing the same river segment. Each county has some authority over decisions that affect water quality, such as enforcement of zoning rules which determine how close to rivers informal housing units (e.g. *favellas*) without adequate access to water or sewage networks can locate, or the selective enforcement of pollution control standards at emitting factories. The splitting of counties therefore leads to de facto decentralization in the sense that an increasing number of jurisdictions gain control over water quality of more and more localized river segments.

This paper examines the determinants of surface water quality, which is a public resource of fundamental importance. Over one billion people in the world lack sufficient potable water, and 80 percent of sewage and 70 percent of industrial wastes are dumped into surface water untreated (World Water Assessment Program, 2006, 2009; Wardlaw et al. 2010). Diarrhea, whose incidence is related to the lack of access to clean water, kills 1.5 million children every year and accounts for 17 percent of under-5 mortality (WHO 2004; UNICEF & WHO 2009).

Our findings are related to studies of border effects on water and other pollution (Burgess et al 2011, Gray and Shadbegian 2004, Helland and Whitford 2003, Fredriksson and Millimet 2002; Sigman 2005, Sigman 2002, Konisky & Woods 2010). Also related is a literature on groundwater depletion (Foster and Sekhri 2008, Sekhri 2011). More broadly, we contribute to a literature on the determinants of water quality and access to clean water (Kremer et al. 2011a, Kremer et al. 2011b, Miguel & Gugerty 2005, Chattopadhyay & Duflo 2004, Bennear & Olmstead 2008, Choe et al. 1996,, Galiani et al. 2005, Devoto et al. 2011, Olmstead 2010). Finally, our use of border changes as a source of identification is related to a literature on the economic effects of shifts in borders (Redding & Sturm 2008, Chakrabarti & Roy 2007, Wolf 2007, Davis & Weinstein 2002, Holmes 1998).

This policy issue – that decentralization without coordination promotes inefficient resource use – is not unique to water management, and is relevant to the management of any publicly provided good with spillovers. For example, local governments may under-invest in health programs if the positive spillover benefits of improvements in health status (e.g. Miguel and Kremer 2004) to those residing outside the jurisdiction are not taken into account. They may under-invest in training farmers if the economic benefits to the neighbors of the farmers employing new technologies are not incorporated in the cost-benefit analysis of a project (Foster and Rosenzweig 1995).

The rest of the paper is organized as follows. Section 2 provides institutional context. Section 3 develops a theory of inter-jurisdictional externalities in river water pollution and develops predictions for the empirical analysis. Section 4 describes our data, and section 5 presents empirical results. Section 6 offers concluding remarks.

2. Literature and Context

2.1 Academic and Policy Literature on Decentralization

Efforts to improve service delivery through decentralization have received increased attention over the past two decades. A large percentage of World Bank loans are focused on decentralization of utilities and government services and a growing number of World Bank loans are made to sub-national governments (Litvack et al., 1998). The United Nations Development Program (UNDP) has also promoted large decentralization programs in Guatemala, the Philippines, Mali, Thailand, and Uganda as part of governance reform packages (Altmann et al., 2000). However, the relative merits of decentralized versus centralized organization of public services remains a debated topic in the scholarly literature. At issue is balancing the objective of improving accountability and responsiveness of the public sector and the difficulty of providing public goods whose costs or benefits cross jurisdictional boundaries.

The economics literature suggests that when preferences or costs are heterogeneous across localities, decentralization fosters greater efficiency (Oates 1972, List and Mason 2001), improved targeting of services (Foster and Rosenzweig 2002), lower levels of corruption and elite capture (Bardhan and Mookherjee 2000) and increased cooperation in homogenous communities (Bardhan 2000). Conversely, the costs of decentralization include inefficiencies associated with inter-jurisdictional spillovers from linked budgets (Besley and Coate 2000), cross-border pollution, duplication in management (Oates 2001), local management inefficiency (de Janvry, Finan, Sadoulet, 2010), and a "race to the bottom" in policy choices (Cumberland 1981).

Identifying conditions under which decentralization improves the efficiency of the public sector remains a key policy challenge. The policymaking community has noted the relative paucity of empirical evidence for the various arguments for and against decentralization (World Development Report 2000). This lack of empirical evidence is in part due to the difficulty of accurately measuring spillover effects, and in part is a result of the impossibility of isolating the effect of decentralization when it is combined with other legislative reforms. Our empirical strategy seeks to fill this gap.

2.2 Rivers in Brazil

Brazil is a particularly important country in which to study water pollution spillovers because it controls 15% of the world's total fresh water resources (World Resources Institute 2003). Lack of sewage treatment is the most important source of water pollution across the densely populated areas of Brazil. Approximately 18 percent of counties report the presence of open sewers that flow directly into major water systems. Farm runoff is the most important cause of water pollution in rural areas. According to the county environmental census, industrial dumping is also a significant concern in approximately 10 percent of counties.

2.3. Can Counties Affect Water Quality?

Although general environmental policy setting and enforcement is determined at the national and state levels, counties in Brazil have important powers over practices affecting the environment within their jurisdiction. Federal law establishes guidelines, norms, and minimum standards of environmental policy, but the importance of county government participation in environmental policymaking has been acknowledged by both state and federal law. The Federal Constitution (1988) allows counties to pass laws establishing local environmental standards that are more strict than the state/federal norms (Engenharia and Projetos 2006), and to enforce standards within their jurisdiction.

Counties are able to fine and tax their community members for activities that cause pollution. In addition, they are able to forbid highly polluting practices and use zoning regulations to reduce direct runoff. The use of these enforcement mechanisms may not be evenly distributed within a county: the county administration has an incentive to increase spending on enforcement of pollution restrictions in areas of the county where pollution will be most harmful to community members. Our conversations with water management practitioners in Brazil suggest that the most important way counties affect water quality is through the application of zoning policies. Local policymakers determine whether or not to permit temporary housing with inadequate infrastructural support close to water sources.[4] The National Water Agency (ANA) in Brazil has recognized the associated pollution externality problems, and inter-county cooperation has become a focus of water management policy in Brazil in recent years (Brannstrom 2004; Formiga-Johnsson and Kemper 2005).

Ferraz (2007) documents another way in which counties influence state regulatory agencies. He finds a significant increase in the number of environmental licenses awarded in years of county mayoral elections in the state of Sao Paulo, as local politicians pressure bureaucrats to allow more polluting firms to operate.

---

[4] County governments also play an important role in extending sanitation services in peripheral regions that lack access to the sewer network (Seroa da Motta, 2006).

2.4. The Process of Creating New Counties

Our empirical strategy relies on changes in border crossings induced by the creation of new counties, and in this sub-section we describe the conditions under which new counties are created. Each county (or *municipio*) is composed of a few districts, and prior to the county mayoral election (which occurs every 4 years during our sample period, beginning in 1988), some districts within a county may choose to split off and create a new county. The process of creating new counties begins with a feasibility study on the projected solvency of the potential county and a referendum on the proposal in both halves of the split county. The referenda are followed by a state law passed by the state legislature and signed by the governor (Tomio 2002).

Brazilian counties receive fiscal transfers from both the federal and the state governments, which introduce a monetary incentive to create new counties. In addition to a portion of the income and industrial taxes collected in their jurisdiction, counties receive a population-based transfer called the Municipalities' Participation Fund (FPM) which, due to its allocation formula, rewards smaller counties. County splits can also occur due to disagreements over the amount of municipal funds used in the various districts of the original county, differences in economic activity across districts, or the large size of the original county (Bremaeker 1992). Splits can also occur for purely administrative reasons, or in order to better represent the political affiliation of the district that leaves the original county (de Noronha, 1995). To the extent that counties have policymaking authority over any publicly provided good, the creation of new counties is a form of decentralization in the delivery of that public good (e.g. two smaller governments rather than one larger one are supplying the service to the same population).

The process of choosing counties to re-district is not random (Weese 2011), and not necessarily uncorrelated with variables that affect changes in water quality. For example, if districts with large increases in population density are more likely to separate from the county, then changes in boundary crossings would be spuriously correlated with changes in water quality. Given the specific pattern of pollution spillovers we report (pollution increases at an increasing rate as the river nears the exit border and pollution decreases after it crosses the border), such observed variables would have to take a very specific (and unlikely) form to be a concern for our interpretation of the results. We theoretically derive the implications of endogeneity in appendix 2, and show that the shape of the pollution function we estimate cannot be easily explained by this form of endogeneity. Furthermore, we add location-specific (i.e. monitoring station) trends to account for changes in population density and economic activity that are gradual, and identify the spillover effects only from the changes in county borders that occur at precise election dates every 4 years.

3. A Model of Pollution Externalities on a River

We model a river on a unit line flowing from left to right, with population distributed along the river according to a probability density function $f(x)$ (see Figure 1). The production and consumption is location-specific: each person at location x consumes q_x and gains utility $u(q_x)$. There is a one-to-one relationship between this consumption and the pollution this person emits into the river: the consumption of q_x generates q_x units of pollution. This setup most accurately describes municipal pollution from population settlements located close to rivers, which is the most relevant setting for our empirical implementation. County governments have the largest influence on such municipal pollution through their zoning practices and expenditures on sewage treatment, and the pollution measure used in the empirical analysis (Biochemical Oxygen Demand or BOD) reflects emissions from municipal point sources.

Pollution emitted at location *x* adversely affects people located downstream of *x*. This pollution exponentially decays as the river flows, and thus the pollution "felt" at downstream point *t* of the emission q_x is $q_x \cdot e^{-(t-x)}$.[5] A social planner decides how much consumption (and pollution) to allow at each location within her jurisdiction by trading off the utility of consumption against the welfare cost of the pollution downstream, subject to the constraint that pollution at any point *x* does not exceed $\bar{q}$, some natural limit on the 'need to pollute'.[6] We first analyze the case where the entire river falls under one jurisdiction, and then examine the effects of a jurisdictional split on the pollution function.

A. Pollution Prior to a Jurisdictional Split

At each point *x* the social planner chooses q_x to maximize the utility that the mass of individuals at *x* receive from consuming q_x net of the harm the associated pollution causes downstream:

$$W = f(x) \cdot u(q_x) - \int_x^1 q_x \cdot e^{-(t-x)} \cdot f(t) \cdot dt \quad \text{(subject to } q_x \leq \bar{q}\text{)} \qquad (1)$$

This yields the first-order condition:

$$f(x) \cdot u'(q_x) = \int_x^1 e^{-(t-x)} \cdot f(t) \cdot dt \quad + \lambda \qquad (2)$$

where λ is the shadow value of the $\bar{q}$ constraint.

---

[5] This is a reasonable decay function for our preferred measure of pollution used in the empirical analysis: Biochemical Oxygen Demand (BOD). The rate of deoxygenation in rivers is commonly modeled using an exponential decay rate. See, for instance, the Streeter-Phelps model (Tchobanoglous and Schroeder, 1987). This specific functional form for decay also makes the analytical solution for optimal emissions tractable, but is not crucial to predictions that we will take to the data.

[6] The $\overline{q}$ constraint is added to the model only for convenience – to avoid arbitrarily large pollution at the edge of the river. As discussed below, $\overline{q}$ does not play a crucial role in generating the empirical predictions.

We now provide an analytical solution for optimal emissions for the simple case of the uniformly distributed population of mass 1 and log utility [$u(q_x) = \ln(q_x)$], and then numerically simulate a pollution function implied by those emissions. The optimal per-person pollution allowance at point *x* is $q_x^* = \min(\frac{1}{1-e^{-(1-x)}}, \overline{q})$.[7] The solid blue line in Figure 2 plots q_x^* for a $\bar{q}$ value of 40. Pollution and consumption allowances increase to the right, since the harm caused by upstream emissions is greater than the harm caused by emissions close to the exiting border out of the jurisdiction.

The actual pollution level felt at any point *y* on the river is the accumulation of all (decayed) pollution allowances to the left of point *y*: $P(y) = \int_0^y q_x^* e^{-(y-x)} \cdot f(x) \cdot dx$ (3)

We numerically integrate and plot *P(y)* in Figure 3 in the absence of a simple analytical solution for this integral. The figure shows that pollution level in a river increases as we head towards a river's exit point out of the jurisdiction, due to the county's strategic optimizing behavior to limit harm to its own constituents.

B. Effect of a Jurisdictional Split on Pollution

The empirical analysis will examine changes in strategic behavior upstream once counties split to create additional jurisdictions. To theoretically predict the effect of a jurisdictional split on water pollution in our model, we introduce a county split at location 0.5 on the river and solve for q_x^u and $P^u(y)$ for the upstream (*u*) county, which is now only concerned about the harm its consumption decisions cause to its own constituents located in

---

[7] q_x^* switches from $\frac{1}{1-e^{-(1-x)}}$ to $\overline{q}$ at the point $\overline{x} = 1 + \ln(1 - \frac{1}{\overline{q}})$, to the right of which $\frac{1}{1-e^{-(1-x)}}$ gets too large. All our numerical simulations assume $\overline{q}$ =40, and at this value $\overline{x}$ is very close to the river's exit point out of the jurisdiction, so $\overline{q}$ does not play a numerically important role.

the interval [0,0.5], and for q_x^d and $P^d(y)$ for the downstream (*d*) county, which is concerned about its own constituents at [0.5,1].

The dashed line in Figure 3 shows that residents of the upstream half of the county are allowed to consume and pollute much more after the split, but the split causes no change in downstream county residents' pollution. The upstream county allows its residents to pollute more since part of the harm caused by the pollution is now an externality on the downstream county that does not enter its own optimizing calculus. The downstream county experiences no such change in the tradeoff between utility and perceived harm. Figure 3 shows that overall pollution level in the river increases due to these negative spillovers brought about by the county split, and that downstream county residents are far worse off. The pollution function is no longer monotonically increasing since there is a sharp discontinuity in the consumption-pollution tradeoff calculus for the two social planners making decisions immediately to the left and to the right of the split.

C. Summary of Testable Predictions

The shapes of the pollution functions in Figure 3 summarize the four main predictions from this simple theory of inter-jurisdictional pollution externalities that we will take to the data:

Prediction 1: Pollution increases as the river travels towards the downstream border where it will exit the county (the *positive* slope of the pollution function in Figure 3).

Prediction 2: Pollution increases at an increasing rate as river heads closer and closer to the downstream border (the *increasing* slope of the pollution function in Figure 3).

Prediction 3: The discontinuous drop in new emissions at the border causes an inflection in the pollution function there. The downstream county does not allow as much new emissions in the most upstream part of its own jurisdiction, and thus the accumulated pollution does not continue to increase as quickly across the border.[8] This creates the *structural break in the slope* of the pollution function in Figure 3.

<u>Prediction 4</u>: There is a larger increase in pollution between an upstream and a downstream point if the river segment defined by those points crosses more borders. This reflects the level effect in Figure 3 comparing the solid orange and dashed blue lines (i.e. before and after the jurisdictional split).

These four predictions jointly determine the shape of the pollution function as the river travels downstream and crosses borders. We use pollution monitoring data to test all four predictions, and in turn estimate the shape of the entire pollution function.

D. Endogenous County Splitting in Densely Populated Areas

A key concern with our estimation strategy to identify the effect of decentralization on pollution spillovers is the possibility of endogenous splitting of jurisdictions in areas with increasing population density or economic activity (where pollution problems are worsening for an independent reason). We derive the implications of this type of endogeneity on the shape of the pollution function in Appendix B. To do so, we assume that there is a second source of pollution related to population density or economic activity that the county government does not control, but which is observationally indistinguishable from the emissions controlled by the county.

Appendix B shows that if county splits occur in high density or economically active areas, the pollution function should continue to display a positive slope downstream of the border. The intuition for this result is that in high density areas, the 'endogenous' source of emissions related to economic activity remains high in the upstream portion of the second county, which makes it more unlikely for us to observe the structural break in the pollution

---

[8] As the river travels further away from the border, new emissions in the downstream county increase. This implies that the pollution function should also exhibit non-linearity downstream of the border.

trend at the border. A stronger version of the statement in Prediction 3 from the theory (that the border causes an inflection point in the pollution function in Figure 3) thus provides an empirical test of the possibility that our regressions merely pick up fluctuations in pollution caused by population density or economic activity changes that have nothing to do with strategic spillovers at borders. Observing a break in the pollution trend at the border, and especially the pollution function switching from positive to negative slope at the border would be stronger evidence in favor of strategic spillovers.

4. Data and Identification Strategy

4.1. An Example of our Identification Strategy

Figure 4 presents a sample map of the evolution of county boundaries from the state of Rio de Janeiro which illustrates our basic identification strategy. Points S1 and S2 mark the locations of two water quality monitoring stations on the river segment flowing from S1 to S2. The theory suggests that the following three variables are important determinants of pollution on this segment: the location of station 1 relative to the nearest downstream border (from S1 to B, which will be termed distance **1D** in Figure 5 and in our empirical analysis), the location of station 2 relative to the nearest upstream border (from B to S2, termed distance **U2** in Figure 5), and the number of county boundary crossings (1, at point B).[9]

The theoretical model predicts that the pollution level at station 1 is expected to be higher when 1D is smaller, or in other words, the closer station 1 is to the downstream border (prediction 1). The change in pollution from S1 to S2 should be greater when more county boundaries are crossed (prediction 4). The inflection point in the pollution function

---

[9] In our empirical work, we will also control for the location of station 1 relative to the nearest upstream border (distance U1, which is the segment from from X to S1), and the distance of station 2 from the nearest downstream border (distance 2D, which is the segment S2 to Y).

at the border (prediction 3) implies that the effect of distance **U2** on the pollution level measured at S2 should be smaller than the effect of distance **1D**.

It is difficult to empirically identify spillover effects using cross-sectional variation in these variables, because for two different river segments of similar length the number of border crossings and distances of monitoring stations to borders would be correlated with average county size and other county characteristics in those regions. Geographic and hydrologic differences can affect the attenuation rate for pollution (Stream Solute Workshop, 1990), and economic and demographic differences can affect the levels of pollution across different rivers. Our empirical strategy involves adding station pair (S1,S2) fixed effects, and examining the effects of changes in distances and in border crossing on the inter-temporal changes in water quality at S1 and S2.

The bottom panel of Figure 4 demonstrates how border crossings and distances to borders vary over time for the same pair of monitoring stations even when the locations of those stations remain fixed. Both Barra Mansa county and adjacent Resende county experienced splits between 1991 and 2001, which allows us to provide a concise illustration of variables 1D and U2. The counties outlined in red were recognized as separate counties by state law after the 1994 and 1998 elections respectively. The creation of Porto Real from Resende county caused distance **1D** (from station 1 to the nearest downstream exit border) to decrease from S1-B to S1-A. Prior to 1994 the Resende leadership was trading off the benefits of pollution allowance around S1 against the costs of pollution to all downstream constituents located along segment S1-B. After 1994 some of those downstream users were no longer Resende voters, and thus the political calculus that determined pollution allowances at S1 changed. Our regressions with the river segment fixed effects identify the inter-temporal *change in* pollution measured at S1 as a result of the *change in* S1's distance to the nearest exit border.

The number of border crossings for the river segment S1-S2 increased from 1 to 3 when Porto Real was created. Prediction 4 implies that since the two new counties now have greater incentives to pollute just upstream of their respective exit borders (i.e. close to points A and B), we should observe that after the split, water quality deteriorates more as the river flows from S1 to S2.

Finally, the creation of county Quatis through the other split reduced distance **U2** of station 2 from the nearest entering border (B-S2 to C-S2). This second type of split allows us to identify the effect of U2 (i.e. the slope downstream of the border) while controlling for the station pair fixed effects.

4.2. Data

The empirical analysis uses water quality measures taken at 372 station pairs across Brazil (see Figures 6 and 7) sampled at quarterly intervals between 1990 and 2007, which results in an unbalanced panel of 8,878 individual biochemical oxygen demand (BOD) observations, or 5,989 unique observations for "station pairs" when stations are matched to the neighboring station along the river. We convert our data to the station-pair format for some of the analysis, since a pair of adjacent stations defines the river segment over which both number of border crossings and pollution changes can be measured.

The preferred pollution variable (BOD) is a measure of the milligrams of oxygen used over a five day period in oxidizing the organic matter contained in one liter of river water (Tchobanoglous and Schroeder, 1985). Higher BOD is associated with increased bacterial count and organisms in the water, which accumulate wherever there is a high level of pollution from organic matter. It is commonly used to measure pollution from industrial, sewage, and runoff sources, and indicates the general health of the river. BOD is the preferred measure used in the prior literature (e.g. Sigman 2002) because it is relatively easily measured by standard procedures, helping to ensure data quality and consistency. BOD tends to be measurable farther downstream than some other pollutants, which makes it appropriate for a study on inter-jurisdictional spillovers. More details about the properties of BOD are in Appendix B.

Table 1 shows that BOD concentrations in Brazilian rivers are relatively high on average, but with large variation. Unpolluted water typically has BOD with levels of 2 or less, while moderately polluted water has BOD levels of between 4 and 8. Treated sewage has BOD levels of 20, while raw sewage has BOD levels of 300 or more (Tebbutt, 1992). Water in our sample has a median (mean) concentration of approximately 2 (3.5). Water quality is worse on average and has higher variance near county boundaries: average BOD levels at stations within 3 kilometers of a county border are nearly one milligram per liter higher than at stations more than 3 kilometers away from any border. This provides some descriptive indication of the basic premise of the theory of cross-border externalities.

Using Geographic Information Systems (GIS) modeling, we combine water quality monitoring station data with maps of rivers to measure changes in BOD as the river flows from an upstream station (S1) to a downstream station (S2). We use elevation maps to identify the direction of water flow. We overlay maps of counties (or *municipios*) and catalog the number of jurisdictional boundaries the river crosses (see Figure 6), distances traversed in each jurisdiction, and distances of monitoring stations to nearest upstream and downstream borders in that year.[10]

---

[10] The county boundaries, and therefore the distances to borders, change over time. These distance calculations are subject to some measurement error, since GIS maps of county boundaries exhibit small variations from year to year. This measurement error likely leads to some attenuation bias.

Brazil re-drew its county boundaries four times between 1990 and 2007, which implies that each water quality observation for a station falls into one of five different county boundary regimes. The number of counties in Brazil increased from 4492 in 1991 to 5807 in 2005. 191 of the 372 station-pairs in the sample experienced at least one border change during the sample period. 3252 of the 5989 water quality observations (i.e. 55%) are for those 191 station-pairs. Each river segment between a pair of stations traverses through four counties and crosses eight borders on average. Monitoring stations in our sample are on average 6.7 kilometers downstream from the last county border, and 9 kilometers from the river's exit point out of the county.

5. Estimation Strategy and Results

5.1 Station Level Regressions

To analyze the effects of border crossings along with the three other predictions of the model, we will ultimately set up our data in a 'station-pair' format where each unit of observation is an upstream-downstream pair of adjacent stations and the dependent variable measures the change in BOD from the upstream station 1 to the downstream station 2. However, we begin with a simpler station-level specification to test the first three theoretical predictions, because regression coefficients are easier to interpret in this format. The estimating equation at the station level is:

$$\begin{aligned}\ln(BOD_{i,t}) = \alpha_i + \sum \delta_{\text{basin-month}} + \sum \gamma_{\text{basin-year}} + \beta_1 \cdot \text{Dist\_Upstr\_Border\_to\_Station } (\mathbf{U1})_{i,t} + \beta_2 \cdot \mathbf{U1}^2_{i.t} \\ + \beta_3 \cdot \text{Dist\_Station\_to\_Downstr\_Border } (\mathbf{1D})_{i,t} + \beta_4 \cdot \mathbf{1D}^2_{i.t} + \sum_k \lambda_k \cdot X^k_{i,t} + \alpha_i \cdot \theta_t + \varepsilon_{i,t}\end{aligned} \tag{4}$$

Refer to Figure 5 for our notation on distances U1 and 1D. This equation examines the determinants of variation in BOD over time, controlling for a station fixed effect, seasonal and annual variations in pollution, and time-varying distances to the nearest upstream (U1) and downstream (1D) borders. We have about 16 water pollution observations for each station on average. We control for station fixed effects (α_i) so that the coefficients on the distance variables of interest (β_1–β_4) are only identified from changes in border locations over time. Another advantage of the location fixed effects estimates is that natural geographic variation in pollutants and the pollution attenuation rate across different rivers would make cross-sectional estimates less precise. To account for the natural seasonal variation in water quality both within and across years and any secular trends in water quality and decentralization, we include year dummies interacted with indicator variables for each of the eight river basins (γ), and also month dummies specific to each river basin δ (since seasonal effects may vary across the river basins).[11] X is a vector time-varying controls for key factors which might affect both water quality and county splits, including GDP, population and area of the county where the monitoring station is located. $\alpha_i \cdot \theta_t$ are trends specific to each station that account for gradual changes in unmeasured components of geography, economic activity or demographics, ensuring that the identification of the distance variables of interest is based only on the sharply dated county mayoral elections and associated county splits.

There are two distance variables of interest: (1) Distance from the upstream border to the pollution monitoring station (U1 in Figure 5), and (2) Distance from the station to the border further downstream which is the river's exit point out of the county (1D).[12] Prediction 1 from the theory states that pollution should increase when 1D decreases. Prediction 2 states that the rate of increase in pollution should be faster as 1D gets smaller.

---

[11] Some specifications in the regression table will add fixed effects for every month of every year, or fixed effects for every month of every year specific to each river basin and demonstrate that the results are robust to these alternative ways of controlling for seasonality and trends.

[12] Using GIS we measure distance along the river, and in most cases this is longer than straight-line crow-fly distance.

To test this prediction we include a non-linear (squared) term for 2D in the specification. Prediction 3 states that the coefficients on 1D and U1 should be different from each other, because this corresponds to a change in slope from upstream to downstream of the border.[13]

5.2 Results with Station-Level Regressions

Table 2 shows the results of estimating equation 4 above. In model (1), the coefficient on the variable **1D** (distance from station to downstream border) shows evidence consistent with prediction 1 from the theory. For every kilometer closer that a river gets to its exit point out of the county, pollution increases by 2%. Furthermore, the statistically significant coefficient on **1D^2** (distance squared) shows evidence consistent with prediction 2: the increase in pollution as the border moves closer occurs at a faster rate the closer we get to the border. Ten kilometers from the border, pollution increases at a rate of 1.3% per kilometer, whereas 1 kilometer from the border it increases at a rate of 1.9%.

Next we examine prediction 3 – that there should be a break in the slope of the pollution function at the border (or at least an inflection) – by statistically comparing the coefficients on U1 and 1D.[14] As predicted by theory, we find that the slope downstream from a border is more negative than the slope upstream of that border, suggesting a slowing of emissions as the river enters a new county. The coefficient on U1 itself is not statistically different from zero, but the slopes implied by two coefficients (+1.996 and -0.250) are statistically different from each other (p-value for F-test = 0.076). Model (2) adds a quadratic term for U1, and the statistical evidence in favor of the break in trend remains strong. The coefficients in model (2) imply that the slope of the pollution function changes

---

[13] Note that we measure pollution at monitoring stations, not at borders, and therefore this test compares the slopes of the pollution function upstream of a monitoring station (which is downstream of a border) and downstream from that monitoring station (which is upstream of the next border).

[14] The appropriate statistical test is coefficient on U1 = - coefficient on 1D, since the coefficient on 1D gives us an indication of the *negative* of the slope of the pollution function.

from +2.0 to -3.3 as we cross the border (p-value ranges from 0.08 to 0.12 depending on the distance from the border at which the F-statistic is evaluated).

The next three specifications show that the evidence in favor of predictions 1, 2 and 3 from the theory are robust to alternative ways of controlling for seasonality and trends and to adding other controls.[15] These results imply that the pollution function we estimate is very similar in shape to the pollution function implied by the simple theory of externalities (Figure 3). Pollution increases at an increasing rate as the river approaches the downstream (exit) border, but the next county does not allow as much new emissions at its own most upstream locations. Furthermore, the pollution function actually has a negative slope just downstream of the border, and appendix 2 shows that this observation would be very difficult to reconcile with the possibility that the spillover effects we document are spuriously generated by new borders being drawn in areas with high (or increasing) density or economic activity.

The last two regressions add a linear trend specific to each station to control for gradual changes in economic activity, geography or population density that may be correlated with county splits. The trends absorb some of the variation in the data, and the standard errors on our coefficients of interest increase. We still find a statistically significant (and slightly larger) effect consistent with Prediction 1: we observe a 2.9-3.1% increase in pollution for every kilometer closer the river gets to the downstream (exit) border. The squared term for distance **1D** remains equally large, at 0.34-0.37, but is now significant at a 10% confidence level. Once we allow for non-linearity in **U1** (which permits a more appropriate test of the structural break *at the border*), the slope of the pollution function on

---

[15] Model 3 introduces fixed effects for every month of every year, and model 4 is more stringent with those fixed effects specific to each river basin. Rivers sometimes form borders and we control for that directly in model 5. Interestingly those (small and statistically insignificant) coefficients imply that when two counties share the river as their border, there is not as much pollution allowed.

either side of the border (implied by coefficients on **1D** and **U1**) are different from each other with a p-value of 0.15.

5.3 Station-Pair Regressions

The simple station-level set up does not allow us to test prediction 4 from the theory, since number of border crossings is not defined for a single station. We now set up our data such that a pair of water pollution observations taken in the same quarter from adjacent water quality monitoring stations on the same river forms our unit of observation. We examine the determinants of increases in water pollution from an upstream point (station 1 in Figure 5) to a downstream point (station 2) as a function of county borders crossed in between, and distances of each station to the nearest upstream and downstream borders:[16]

$$\begin{aligned}\ln(BOD_{i,t}^{2}) - \ln(BOD_{i,t}^{1}) = {}& \alpha_i + \alpha_i \cdot \theta_t + \sum \delta_{\text{basin-month}} + \sum \gamma_{\text{basin-year}} + \beta_1 \cdot \text{Border\_Crossings} \\ & + \beta_2 \cdot \text{Dist\_Station1\_to\_Downstr\_Border } (\mathbf{1D})_{i,t} + \beta_3 \cdot \text{Dist\_Upstr\_Border\_to\_Station2 } (\mathbf{U2})_{i,t} \\ & + \beta_4 \cdot \mathbf{U1}_{i,t} + \beta_5 \cdot \mathbf{2D}_{i.t} + \sum_k \lambda_k \cdot X_{i,t}^{k} + \varepsilon_{i,t}\end{aligned} \tag{5}$$

The three coefficients of interest in this regression corresponding to theoretical predictions are β_1, β_2, and β_3. Prediction 4 states that we should expect β_1 to be positive: pollution increases along a river as a larger number of county borders are crossed. Prediction 1 implies that β_2 is *positive*: as station 1's distance to the exit point out of the county (1D) decreases, pollution measured at station 1 [ln(BOD^1)] should increase, which in turn makes the dependent variable ΔBOD =[ln(BOD^2) – ln(BOD^1)] smaller.[17] Prediction 3 implies that the β_2 and β_3 coefficients should not be equal, and in particular, a negative β_3 would indicate that the pollution function actually changes slope from positive to negative across the border.

---

[16] Station pairs not separated by a county border are categorized as having zero distances to the borders that don't exist, and we include indicator variables for these cases in all regressions.

[17] Note that this is opposite to the expected sign on variable 1D in Table 1, because now pollution measured at station 1 [ln(BOD^1)] appears on the left-hand-side of the equation with a *negative* sign.

In the station pair setup we have a larger set of variables in X, since each time-varying control – GDP, population, area is now measured separately for the county where the upstream monitoring station (1) is located, the county where the downstream station (2) is located, and the distance-weighted average for the "intermediate" counties that the river segment flows through while traveling from station 1 to station 2. Furthermore, in this specification and in all subsequent specifications we control for linear trends specific to each station pair (denoted $\alpha_i \cdot \theta_t$ in the estimating equation). Additionally, we control for the distances of each station to the nearest borders that lie outside of the river segment (U1 and 2D in Figure 5) because theory suggests that those variables may affect pollution measured at either station 1 or 2. Other fixed effects are the same as in the station-level analysis.

5.4 Results in Station-Pair Framework

The first specification examines the effects of county border crossings on pollution in the river. For each additional border crossed, pollution increases by 3.1%. The next three specifications show that this finding is robust to including a variety of distance controls, including controls for the river forming the county border at various locations. We find robust evidence in favor of prediction 4 in the model: that pollution increases as more borders are crossed. By including a station-pair fixed effect, we derive this evidence on the basis of additional borders crossed by the same river segment in later years due to county splitting. We directly control for time-varying GDP and population in all counties[18], and add linear trends specific to each river segment, to account for unobserved components of population density and economic activity which may be growing gradually over time and which may be correlated with county splitting.

---

[18] GDP and population affect pollution in the expected direction: population in the county where station 2 is located increases the dependent variable $\ln(BOD^2) - \ln(BOD^1)$], and GDP in the county where station 1 is located decreases the dependent variable [$\ln(BOD^2) - \ln(BOD^1)$].

Re-examining the effects of distances to borders in the station pair setup, we again find strong evidence in favor of the model's Prediction 1: as the river approaches the downstream exit border out of the county (distance **1D**), pollution increases by 1.5-1.7% per kilometer. The coefficient on **U2** shows that as the river crosses over to the other side of the border, the slope of the pollution function reverses form +1.5% to -3.1% per kilometer. The structural break in slope is statistically significant: the p-value of the difference in 1D and U2 coefficients is 0.077. The distance results are therefore consistent with both predictions 1 and 3 from the model: pollution increases as you head to the downstream (exit) border, but there is a structural break in the slope of the pollution function at the border, and the second (downstream) county appears to not allow new emissions at the same rate in its most upstream part.

Table 3 indicates that the evidence in favor of predictions 1, 3 and 4 is quite robust to alternate specifications and controls. First, we always control for additional distance variables for each station to their nearest borders outside of the river segment (U1 and 2D in Figure 5). Second, we add controls for distances that the river forms a border in different parts of the river segment. Third, we always control for GDP, population and area of the counties where the two stations are located, as well as the counties that fall in between.

5.5 Flexible Non-linearity Tests of Predictions 2 and 3

In Table 4, we examine (a) the non-linearity in the pollution function at varying distances from the downstream exit border (prediction 2), without imposing a quadratic parametric restriction on the shape of the function like we did in Table 2, and (b) the structural break in the slope of the pollution function at the border, evaluated with data from monitoring stations at varying distances from the border.

For prediction 2, we study the effects of variable **1D** (distance of station 1 to nearest downstream border) separately for stations located close to the border (i.e. within 1, 3, 5, or 15 kilometers from the border in the different specifications) and stations located farther from the border. The theoretical expectation is that the effect of 1D measured close to the border should be larger than the effect measured farther away from the border. In all cases, estimated standard errors are large, and we cannot precisely identify effects using these bins at varying distances. The point estimates on coefficients imply a pattern consistent with prediction 2: that increases in pollution per kilometer are the largest very, very close to the border. However, the large standard errors imply that we cannot reject linearity.

For prediction 3 (inflection or break in slope at the border) we test whether the coefficients on 1D and U2 are different from each other, when they are evaluated within 1km of the border (and then 3, 5 and 15km in turn). The coefficients are all consistent with the theoretical prediction, but they are often imprecise. Evaluated within 1km of the border, the slope changes from +9.4% to -9.5% across the border, but both estimates are imprecise. Evaluated within 3km of the border, the slope changes from +7.8% to -13%, and these coefficients are significantly different from each other. Further away from the border, the estimates are smaller and statistically insignificant.

The coefficient estimates are all generally consistent with the model, but they provide only suggestive evidence. The large standard errors make it clear that the variation in the data is not rich enough to support the border changes based identification of pollution effects when the data are parsed into small bins.

6. Mechanisms, Alternative Explanations and Robustness of Empirical Results

This section reports the results of a few tests to check for robustness of the effects we report, to examine the underlying mechanisms by which these effects are realized, and to consider plausible alternative explanations for the empirical patterns we document.

6.1 Robustness Checks

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 5 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 $\mathbf{1D}^2$ 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 6 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.

6.2 Mechanisms

The specific spatial patterns of pollution within municipios that we document, and our simple political economy based interpretation of those patterns, require that politicians have access to policy levers that allow them to spatially target emissions into rivers. Our conversations with water management practitioners in Brazil indicate that there are two primary ways that this could happen: (a) zoning regulations that determine how close to water bodies slums of *favellas* with inadequate water and sanitation support are allowed to locate, and (b) differential enforcement of pollution permit regulations for firms and settlements in downstream locations relative to portions of the county further upstream. This sub-section uses additional data sources to examine these mechanisms.

Table 7 uses the 2002 Environmental Census conducted by the Brazilian National Statistical Agency to describe the major sources of water pollution. The single most common source of pollution reported by counties is domestic sewage. 12% of all counties (and 35% of those with water pollution problems) cite poor enforcement of pollution regulations. Panel B in Table 7 shows that counties engage in a wide range of activities to address pollution problems, including monitoring and fining polluters. Feler and Henderson (2011) report many instances of systematic under-provision of public water and sanitation services in slums across Brazil, strong prevalence of informal housing, and they document strategic under-provision of public services by counties.

It is easier for policymakers to spatially target emissions for point-source pollutants than for diffuse sources of pollutants such as run-off from farms. If the political economy mechanisms above are correct, then we should be less likely to observe the same spatial patterns of pollution close to and across borders when we examine pollution measures associated with more diffuse sources of emissions. The preferred water pollution measure we use (Biochemical Oxygen Demand) is most closely associated with municipal sewage, although it may also indicate chemical pollution from firms or farms. In Table 8 we examine the effects of borders on three pollution measures associated with soil erosion: suspended solids, turbidity, and electrical conductivity of the water. Erosion tends to be heavily influenced by non-point source (agricultural) pollution. In these "placebo" tests, we do not find any strong evidence of strategic pollution shifting around borders. In other words, for emissions sources that are not easy for county policymakers to manipulate, there are no systematic effects of county borders.

6.3 Alternative Explanations for our Results

A key concern for our county re-districting based estimation strategy is whether the non-random choice of locations where new borders are drawn can explain the patterns in the data we report. The most straightforward example is that if districts with increases in population density or economic activity are more likely to separate from the county, then water quality may deteriorate around new borders anyway. In addition to controlling for time-varying population density and local GDP, and for trends specific to each location to account for gradual changes in other unmeasured attributes of each location, a theoretical exercise in Appendix B shows that it is not easy to generate all four results that we document in the data (associated with predictions 1-4 from the theory) on the basis of alternative mechanisms. Appendix B considers the specific case of new borders drawn in high density areas and shows that the result associated with prediction 3 (the structural break in slope at the border) is not replicated in that model. Another potential theory – that new borders are drawn in outlying low density areas of counties in between two areas of increasing density - is not easy to reconcile with the empirical results associated with predictions 1 and 2 (that pollution increases at an increasing rate as we approach borders).

It is still important to be precise about all the socio-economic changes that occur when counties split, in order to fully understand and correctly interpret the effect of additional border crossings (prediction 4). First, the total budget allocated to the local governments increase when counties split, because the combined replacement budget of the two new counties typically exceeds the budget of the original one large county. Second, the socio-economic or demographic composition of each new county may be different from the original county.

We can examine the likely budgetary impacts of county splitting on water quality by running a fixed effects regression comparing budgets before and after the split. After splits, the new smaller counties see the county health and sanitation spending increase by R$13.2 per person over the spending in the larger county that they were a part of in the previous year. For the average county in Brazil, this translates into a 20% increase in expenditures. The deteriorating water quality we observe after splits occurs *in spite of* the increase in local spending that is likely to have a cleansing effect on water. Thus pollution externalities across borders are possibly even larger than those implied by our estimates.

After a split, new smaller counties may be more homogenous than the original larger county, which in itself may reallocate resources towards a variety of public goods including pollution abatement (Alesina, Baqir and Easterly, 1999). Again, this would offset the spillover effects we document, and this mechanism makes it more difficult for us to find evidence of pollution externalities at borders.

An example of yet another type of concern would be that counties with strong leadership or community involvement across districts are less likely to have districts separating, so that water quality would in general be lower in split areas. Since our regressions control for location fixed effects and inference is based only on *changes in water* *quality over time in the same river segment*, such level differences in water quality are not a bias concern. This may indicate, however, that our empirical identification and data variation are driven by certain types of counties, which limits the applicability of our results to other contexts.

7. Conclusion

In summary, we report a variety of congruent evidence that counties behave strategically in deciding where and how much to pollute, consistent with a simple theory of externalities. In this section we consider two further implications of these results for policy:

a) Are the magnitude of spillovers and the excess pollution induced by strategic behavior large enough to warrant policy intervention?

b) Why have jurisdictions not already negotiated solutions to the externality issue as predicted by the Coase Theorem (Coase 1960)?

In order to tackle the first question, we infer the magnitude of pollution spillovers implied by our estimates. We compute the predicted change in Biochemical Oxygen Demand (BOD) associated with a new county border drawn close to a representative monitoring station in our sample, and then interpret the magnitude of that change in BOD in terms of U.S. Environmental Protection Agency (EPA) water quality standards. County redistricting introduces one new border between the modal monitoring station in the sample and its downstream neighbor, and the new border reduces that station's distance to the nearest downstream exit point out of the county from 6 kilometers to 3 kilometers. Our regression estimates imply that these two changes jointly increase the BOD measured at this station by about 9%. At the sample average station, this translates into an increase in BOD from 3.5 to 3.9 mg/L. While this is an environmentally significant change, in that water with BOD exceeding 4 mg/L is not considered acceptable for recreational use (Sigman 2002), these magnitudes are likely not so large or visible that we would expect political action or policy changes to occur on the basis of it.

To address the second question, we note that Brazil *has* shifted toward integrated river basin management in recent years in response to deteriorating water quality and increased conflicts over water use (Porto and Kelman 2005). A 1997 national water management act encouraged the formation of "river basin committees" with participation by all actors with a vested interest: federal, state and county government representatives, user groups, and members of civil society. River basins are the natural geographic units which encompass the interests of all potential users, and many water sector experts promote management at this level (Saleth 2002, Mody 2004, Abu-Zeid and Biswas 1996). These committees are being introduced slowly across Brazil, and the earliest innovations started occurring towards the end of our sample period. Furthermore, the literature has noted the difficulty of getting leaders from different municipios with different political affiliations to cooperate on water management decisions (Formiga-Johnsson and Kemper 2005).

A case study-based descriptive literature reports that river-basin management has met with mixed success around the globe, but quantitative or empirical evaluations of such initiatives is lacking (Biswas and Tortajada 2001, and Kemper et al. 2007). Given this paper's findings, this would be a natural topic on which to conduct follow-up research, especially as more and more water basin committees become operational in Brazil and elsewhere.

In summary, we document significant strategic polluting behavior around borders, which suggest that institutions promoting decentralization (national governments, United Nations, World Bank) ought to be more vigilant in assessing externality costs of decentralization, especially around border areas. The spillover costs should be evaluated against other benefits of decentralization such as improved targeting of public services.

References

Abers, Rebecca (2007). "Organizing for Governance: Building Collaboration in Brazilian River Basins." World Development. V. 35. N.8.

Abu-Zeid, M., & Biswas, A. (Eds.). (1996). River basin planning and management. Calcutta: Oxford University Press.

Agencia Nacional de Aguas (2002). *The Evolution of the Organization and Implementation of Water Basin Management in Brazil*, Report prepared by ANA, Brasilia for the 'International Conference of Water Basin Agencies, Madrid, 4-6 November, 2002.

Alesina, Alberto, Reza Baquir and William Easterly (1999). "Public Goods and Ethnic Divisions," *Quarterly Journal of Economics,* Vol CXIV Issue 4 p. 1243-1284.

Altmann, Joern, Ledivina Carino, Richard Flaman, Manfred Kulessa, Ingo Schulz, 2000. "The UNDP Role in Decentralization and Local Governance." New York: United Nations Development Programme, Evaluation Office.

Bardhan, Pranab, (2000) "Irrigation and Cooperation: an Empirical Analysis of 48 Irrigation Communities in South India." *Economic Development and Cultural Change* 48(4): 847-865.

Bardhan, Pranab and Dilip Mookherjee (2006). "Decentralization and Accountability in Infrastructure Delivery in Developing Countries." *The Economic Journal* 116(January): 101-127.

Bennear, Lori S. and Sheila Olmstead (2008). "The Impacts of the `Right to Know': Information Disclosure and the Violation of Drinking Water Standards." *Journal of Environmental Economics and Management* 52(2): 117-130.

Besley, Timothy and Stephen Coate, (2003). "Centralized versus Decentralized Provision of Local Public Goods: A Political Economy Analysis" *Journal of Public Economics* 87(12): 2611-2637.

Biswas, Asit K. and Cecilia Tortajada (eds.) (2001). *Integrated River Basin Management: The Latin American Experience*, Oxford University Press, New Delhi.

Brannstrom, Christian, (2004), "Decentralizing Water Resources Management in Brazil," *The European Journal of Development Research,* Vol.16 Issue 1 pp. 214-234.

Bremaeker, Francois (1992). "Os Novos Municipios: Surgimento, problemas e Solucoes." *Revista de Administracao Municipal,* 40 (206) 88-99.

Burgess, R., M. Hansen, B. Olken, P. Potapov, S. Sieber (2011). "The Political Economy of Deforestation in the Tropics," MIT Department of Economics Working Paper.

Chakrabarti, Rajashri and Joydeep Roy (2007). "Effect of Redrawing of Political Boundaries on Voting Patterns: Evidence from State Reorganization in India." *Federal Reserve Bank of New York Staff Reports No. 301.* Sept. 2007.

Chattopadhyay, Raghabendra and Esther Duflo (2004). "Women as Policy Makers: Evidence from a Randomized Policy Experiment in India." *Econometrica* 72(5): 1409-1443.

Choe, KyeongAe, Dale Whittington and Donald T. Lauria (1996). "The Economic Benefits of Surface Water Quality Improvements in Developing Countries: A Case Study of Davao, Philippines" *Land Economics* 72(4): 519-537.

Cumberland, John H. (1981) "Efficiency and Equity in Interregional Environmental Management," *Review of Regional Studies* 2, pp. 1-9.

Davis, Donald R. and David E. Weinstein (2002). "Bones, Bombs and Break Points: The Geography of Economic Activity." *American Economic Review* 92(5): 1269-1289.

De Janvry, Alain, Frederico Finan, and Elisabeth Sadoulet, (2010) "Local Electoral Accountability and Decentralized Program Performance." *Review of Economics and Statistics*, forthcoming.

Devoto, Florencia, Esther Duflo, Pascaline Dupas, William Parienté, Vincent Pons (2011). "Happiness on Tap: Piped Water Adoption in Urban Morocco." *Unpublished manuscript*: April, 2011.

Engenharia & Projetos (2006). " Manual de Saneamento e Proteção Ambiental para os Municípios." Monte Santa de Minas, MG.

Faguet, Jean-Paul (2004). "Does Decentralization Increase Government Responsiveness to Local Needs? Evidence from Bolivia." *Journal of Public Economics.* 88(3-4): 867-893.

Feler, Leo and V. Henderson (2011). "Exclusionary policies in urban development: Under-servicing migrant households in Brazilian Cities," *Journal of Urban Economics* 69: 253-272.

Ferraz, Claudio (2007). "Electoral Politics and Beaurocratic Discretion: Evidence from Environmental Licenses and Local Elections in Brazil." Working Paper.

Formiga Johnsson, Rosa Maria and Karin Erika Kemper (2005). "Institutional and Policy Analysis of River Basin Management: The Jaguaribe River Basin, Ceará, Brazil." *World Bank Policy Research Working Paper Series No. 3649*, June 2005.

Foster, Andrew D. and Mark Rosenzweig (1995). "Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture." *Journal of Political Economy.* 103(6): 1176-209.

Foster, Andrew and Mark Rosenzweig (2002), "Democratization, Decentralization and the Distribution of Local Public Goods in a Poor Rural Economy," mimeo, Brown University.

Foster, Andrew and S. Sekhri (2008). "Water Markets, Local Networks, and Aquifer Depletion," Working Paper, University of Virginia.

Fredriksson, Per G. and David L. Millimet (2002). "Strategic Interaction and the Determination of Environmental Policy across U.S. States." *Journal of Urban Economics* 51: 101-122.

Galiani, Sebastian, Paul Gertler and Ernesto Schargrodsky (2005). "Water for Life: The Impact of the Privatization of Water Services on Child Mortality." *Journal of Political Economy* 113(1): 83-120.

Gray, Wayne B. and Ronald J. Shadbegian (2004). "`Optimal' Pollution Abatement – Whose Benefits Matter, and How Much?" *Journal of Environmental Economics and Management* 47: 510-534.

Helland, Eric and Andrew B. Whitford (2003). "Pollution Incidence and Political Jurisdiction: Evidence from the TRI." *Journal of Environmental Economics and Management* 46: 403-424.

Hobbs, Greg (2005). History of Colorado River Law, Development and use: A Primer and Look Forward. *Conference proceedings: Hard Times on the Colorado River: Drought, Growth and the Future of the Compact.* University of Colorado School of Law: Natural Resources Law Center, June 8-10, 2005.

Holmes, Thomas (1998). "The Effects of State Policies on the Location of Industry: Evidence from State Borders," *Journal of Political Economy* 106 (4): 667-705

IBGE (2002). *Perfil dos Municipios Brasilieros*, Meio Ambiente.

Jongerden, Joost (2010). "Dams and Politics in Turkey: Utilizing Water, Developing Conflict." *Middle East Policy*, XVII(1), 137-43.

Kemper, K., Blomquist, W., & Dinar, A. (Eds.). (2007). Integrated river basin management through decentralization. New York: Springer.

Konisky, David M. and Neal D. Woods (2010). "Exporting Air Pollution: Regulatory Enforcement and Environmental Free Riding in the United States." *Political Research Quarterly* 63(4): 771-82.

Kremer, Michael, Jessica Leino, Edward Miguel and Alix Zwane, (2011) "Spring Cleaning: Rural Water Impacts, Valuation, and Property Rights Institutions" *Quarterly Journal of Economics,* 126(1): 145-205.

Kremer, Michael, Edward Miguel, Sendhil Mullainathan, Clair Null and Alix Zwane (2011). "Social Engineering: Evidence from a Suite of Take-up Experiments in Kenya" Working Paper.

Lemos, Maria Carmen and Joao Lucio Farias de Oliveira (2004), "Can Water Reform Survive Politics? Institutional Change and River Basin Management in Ceara Northeast Brazil," *World Development* Vol. 32 No.12 pp. 2121-2137.

Levinson, Arik, (1997), "A Note on Environmental Federalism: Interpreting some Contradictory Results," *Journal of Environmental Economics and Management* 33, pp.359-366.

List, John and Charles F. Mason (2001), "Optimal Institutional Arrangements for Transboundary Pollutants in a Second-Best World: Evidence from a Differential Game with Symmetric Players," *Journal of Environmental Economics and Management* 42 pp.277-296.

Litvack, Jennie, Junaid Ahmad, and Richard Bird (1998). "Rethinking Decentralization in Developing Countries," World Bank Sector Studies Series, 21491.

Miguel, Edward and Mary Kay Gugerty (2005). "Ethnic Diversity, Social Sanctions, and Public Goods in Kenya." *Journal of Public Economics*: 89(11-12): 2325-2368.

Miguel, Edward and Kremer, M. (2004). "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities" *Econometrica* 72, No. 1.

Millimet, Daniel L. (2003). "Assessing the Empirical Impact of Environmental Federalism." *Journal of Regional Science.* 43(4): 711-733.

Mody, J. (2004). Achieving accountability through decentralization: Lessons for integrated river basin management. World Bank Policy Research Working Paper.

Mustafa, I. (1994), "The Arab-Israeli Conflict over Water Resources", in J. Isaac and H. Shuval (eds.), *Water and Peace in the Middle East*, Amsterdam: Elsevier, 123-33.

Neves, Estela Maria Souza Costa (2006). A política ambiental e os municípios brasileiros. Doctoral Thesis, Universidade Federal Rural do Rio de Janeiro, Rio de Janeiro.

Noronha, Rudolf de and Elizabeth D. Cardoso, (1996). "Emancipacoes Municipais: Como Ficam os Municipios de Origem?" *Revista de Administracao Municipal, 42 (214)* 67-80.

Oates, W.E., (1972). *Fiscal Federalism*, Harcourt, Brace and Jovanovich, New York, 1972.

Oates, Wallace E. (2001). "Fiscal Competition and European Union: Contrasting Perspective." *Regional Science and Urban Economics.* 31(2-3): 133-145.

Oates, W.E. and R.M. Schwab. (1988) "Economic Competition among Jurisdictions: Efficiency Enhancing or Distortion Inducing?" *Journal of Public Economics,* 35 333-354.

Olmstead, Sheila M. (2010). "The Economics of Water Quality." *Review of Environmental Economics and Policy* 4(1): 44-62.

Porto, M., & Kelman, J. (2005). Water resources policy in brazil. Working Paper.

Postel, Sandra. (1999). "Water Wars I: Farms versus Cities and Nature", in *Pillar of Sand: Can the Irrigation Miracle Last?,* Chapter 6, New York and London: W.W. Norton & Co., 110-32, 281-5.

Redding, Steven and Daniel Sturm. (2008). "The Costs of Remoteness, Evidence from German Division and Reunification," *American Economic Review*, 98(5).

Richards, A. and N. Singh (2002). "Inter-state Water Disputes in India: Insitutions and Policies." *International Journal of Water Resources Development.* 18(4) p.611-625.

Saleth, M. (Ed.). (2002). Water resources and economic development. Northhampton, MA: Edward Elgar.

Sekhri, Sheetal (2011). "Federalism versus Regional Control: Implications for Groundwater Resource in India," Working Paper, University of Virginia.

Seroa da Motta, Ronaldo and Ajax Moreira, (2006). "Efficiency and Regulation in the Sanitation sector in Brazil." *Utilities Policy* 14(3) 185-195.

Sigman, Hilary (2002). "International Spillovers and Water Quality in Rivers: Do Countries Free Ride?" *The American Economic Review* 92(4): 1152-59.

Sigman, Hilary (2005). "Transboundary Spillovers and Decentralization of Environmental Policies." *Journal of Environmental Economics and Management.* 50(1): 82-101.

Stream Solute Workshop, (1990). "Concepts and Methods for Assessing solute dynamics in stream ecosystems" Journal of the North American Benthological Society 9(2) p.95-119.

Tchobanoglous, George and Edward D. Schroeder (1985). *Water Quality: Characteristics, Modeling and Modification.* Reading, MA: Addison-Wesley Pub. Co., 1985.

Tebutt, T. H. Y. *Principles of Water Quality Control.* 5th ed. Oxford, U.K., and Woburn, MA: Butterworth-Heinemann, 1998.

Tomio, Fabricio Ricardo de Limas (2002). "The Creation of Municipalities after the 1988 Constitution, " *Revista Brasileira de Ciencias Sociais.* 17: N8 p.61-89.

UNICEF & WHO (2009). *Diarrhoea: Why Children are Still Dying and What Can Be Done.* Switzerland.

Wardlaw, Tessa, Peter Salama, Clarissa Brocklehurst, Mickey Chopra, and Elizabeth Mason. "Diarrhoea: Why Children are Still Dying and What Can Be Done." *The Lancet* 375(9718): 870-72.

Weese, Eric (2011). "Political Mergers as Coalition Formation: Estimates from Japanese Municipal Mergers", Working Paper, Yale University

Wolf, Aaron (1998). "Conflict and Cooperation Along International Waterways." *Water Policy* 1(2), 251-65.

Wolf, Aaron (ed.) (2002). *Conflict Prevention and Resolution in Water Systems*, Edward Elgar, Northampton, MA.

Wolf, Nikolaus (2007). "Endowments vs. Market Potential: What Explains the Relocation of Industry After the Polish Reunification in 1918." *Explorations in Economic History* 44(1): 22-42.

World Bank, (2004). "World Development Report 2004: Making Services Work for Poor People." Washington, D.C.: The International Bank for Reconstruction and Development/The World Bank.

World Health Organization (2004). *The Global Burden of Disease, 2004 Update*, Switzerland.

World Resources Institute, (2003). "Earthtrends Country Profiles." http://earthtrends.wri.org

World Water Assessment Program. (2006). *Water and Industry Fact Sheet.* Retrieved November 8, 2011, from http://www.unesco.org/water/wwap/facts_figures/water_industry.shtml

World Water Assessment Program. (2009). *United Nations World Water Development Report 3: Water in a Changing World.* Paris, France: UNESCO.

Appendix A: Chemical Properties of 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 Schroeder, 1987).

BOD increases as micro-organisms accumulate to degrade organic material. High levels of organic pollution and BOD are associated with river eutrophication.[19] 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 wastes, 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.[21]

Water Type	**BOD Level**
Unpolluted	2
Highly Polluted	10
Treated Sewage	20-100
Raw Sewage	600
Industrial Waste	up to 25,000[22]

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.[23]

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; water velocity increases the oxidation of the organic pollutants in the water—high flow rates cause increased churning and oxygenation.[24]

---

[19] Chapman, 1996, p. 276-278.

[21] Chapman, 1996. P. 122.

[22] Chapman, 1996, p. 88.

[23] Hounslow, 1995. P. 302.

[24] Chapman, 1996, p. 246-276

Appendix B: Implications of Endogenous County Splitting in Densely Populated Areas for the Empirical Analysis and Interpretation

As discussed in section 3, a key concern with our estimation strategy to identify the effect of decentralization on pollution spillovers is the possibility of endogenous splitting of jurisdictions in areas with high population density (where pollution problems are worsening for an independent reason). Under the particular form of welfare maximizing behavior by the county authority that we've assumed in the theory, there is actually no such endogeneity problem since the authority would respond to (say) a doubling of the population by simply halving each person's consumption allowance. With twice the population, each person's emissions cause double the harm, and so the county authority forces its citizens to cut back on consumption. However, to guide a careful empirical strategy we do want to allow for such endogeneity, so we will now assume that each person at location x emits ε_x in addition to the q_x, but that the ε_x emissions are un-monitored and beyond the control of the county authority. Thus we will model 'endogeneity' as follows: when population density increases at a location, counties are likely to split there, but there is also an independent effect on pollution at those locations since the ε component of emissions are now larger there. We have to also relax the assumption of a uniform population distribution in order to effectively model increasing population density.

Imagine that population doubles (from mass 1 to 2), and that *f(x)* now takes the form of a symmetric triangular distribution, so that the largest increase in population density occurs right around 0.5:

$$f(x) = \left| \begin{array}{ll} 4\mathrm{x} & \text{for } 0 \le x \le 0.5 \\ 4(1-\mathrm{x}) & \text{for } 0.5 \le x \le 1 \end{array} \right. \qquad (4)$$

We will examine the effect of a jurisdictional 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 first-order conditions (2) yield the following solutions for pollution allowances in the upstream and downstream counties:

$$q_x^u = \min\left(\frac{1}{4[1+x-1.5e^{-(0.5-x)}]}, \overline{q}\right) \text{ and } q_x^d = \min\left(\frac{1}{4[e^{-(1-x)}-x]}, \overline{q}\right)$$

The dashed lines in Figure 4 plot the associated pollution function which only accounts for 'strategic' q-type pollution but not the unmonitored ε-type pollution. Pollution increases very sharply upto point 0.5 (and this pollution function is steeper than the corresponding one for the uniform distribution of population) because the strategic motives to pollute more and the effects of increasing population density coincide at locations just left of the split. Unlike the uniform distribution case, pollution monotonically decreases downstream of the split since the downstream county, concerned about the welfare of its citizens, allows relatively little new pollution within its border, and the unusually large inflows of pollution from the upstream county decay as the river flows. This particular difference in the shapes of the dashed blue lines in figures 3 and 4 (non-monotonic quadratic for the uniform distribution versus monotonically decreasing pollution downstream of county borders for the triangular distribution) yields a simple empirical test of the basic premise of the endogeneity concern – that county splits occur in areas of high population density. The intuition for the test is that with population-density based splitting, county borders are likely to be located in areas with high population density, so that when we move away downstream of borders, population density decreases, which lowers observed pollution. As we will see in the next section, our data are consistent with the population density based splitting, so the basic premise of this form of endogeneity is borne out.

Figure 4 also plots a "total pollution function" in solid orange, which aggregates the q-type with the unmonitored ε-type pollution. This function corresponds to the pollution that will be observed in the data (since the data is just "total pollution" aggregated across q-type and ε-type). The three panels of the appendix figure vary the assumed levels of ε-type pollution. 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 that larger amounts of "endogeneity" are present in our empirical analysis – that our regressions merely pick up fluctuations in pollution caused by population density changes that have nothing to do with strategic spillovers. The idea is to compare the shapes of the pollution functions under differing degrees of endogeneity to the empirically estimated shape of the pollution function to see whether the estimates based on the data correspond to large or small endogeneity concerns.

As we add larger amounts of 'endogeneity' (i.e. ε-type pollution), the shape of the total pollution function changes: total pollution keeps increasing to the right of the border, replacing the monotonically decreasing function associated with no endogeneity. This is because population density is largest close to the border, and this is where the emissions of per-person ε-type pollution is the greatest. A comparison of the shapes of the solid blue and the dashed orange lines across the three panels of figure 4 yields an empirical test of the quantitative importance of the endogeneity concern. If the correlation between distance to border to pollution is driven by population density rather than true strategic behavior, then the estimated relationship between pollution levels and distance downstream of border should follow a non-linear inverted-U shaped pattern. Observing a negative linear relationship between downstream distance and pollution would be more consistent with evidence of strategic spillovers. The key insight here is that if county splits occur in high density areas, that has implications for the spatial patterns of "endogenous" population density driven pollution around borders. Examining those spatial patterns allows us to make some empirical inferences on the extent to which the correlation is driven by population changes rather than strategic behavior by counties in the presence of spillovers. We allow for non-linear effects of downstream distance in our empirical work, and always find distance traversed downstream has a linear negative effect. The empirical results reported in this paper are thus likely evidence of strategic behavior as opposed to spurious correlation due to changing density.

Table 1: Summary Statistics

	Obs	Mean	Std. Dev.
Border Variables			
Number of border crossings	5989	8.422	11.218
Distance of station 1 from its nearest upstream border	5989	4.701	8.289
Distance from station 1 to its downstream border	5989	5.759	7.666
Distance of Station 2 from its nearest upstream border	5989	9.634	13.856
Distance from Station 2 to its downstream border	5989	12.358	18.731
County Characteristics			
GDP in Downstream County in 100,000 Reais	5989	4.388	17.343
Population in Downstream County in 100,000 people	5989	0.648	2.023
Area in Downstream County in 100,000 square kilometers	5989	0.025	0.040
GDP in upstream county in 100,000 Reis	5989	3.897	16.629
Population in Upstream County in 100,000 people	5989	0.559	1.721
Area in Upstream county in 100,000 square kilometers	5989	0.020	0.022
Water Quality			
BOD at upstream station	5989	3.529	12.529
BOD at downstream station	5989	3.382	15.247
Water Quality Nearer to and Farther Away from Borders			
Upstream Station BOD			
Within 3 km from downstream border	3162	3.913	15.784
Beyond 3 km from downstream border	2827	3.099	7.311
Downstream BOD			
Within 3 km of upstream border	2407	3.628	20.300
Beyond 3 km of upstream border	3565	3.219	10.603

Table 2: Station Level Regression: Determinants of Changes in Pollution at a Station

Dependent Variable: 100*LogBOD in the upstream station							
Distance from station 1 to Downstream border (1D)	-1.996***	-2.007***	-1.654**	-2.016**	-1.978**	-2.902**	-3.141**
	(0.729)	(0.758)	(0.684)	(0.853)	(0.766)	(1.234)	(1.267)
Squared Distance from station 1 to Downstream border ($1D^2$)	0.033***	0.033***	0.029***	0.033***	0.032***	0.034*	0.037*
	(0.011)	(0.011)	(0.009)	(0.011)	(0.011)	(0.019)	(0.020)
Distance from Upstream border to station 1 (U1)	-0.250	-3.316	-3.618	-3.337	-3.416	0.983	-5.024
	(0.926)	(3.163)	(2.784)	(3.228)	(3.166)	(1.946)	(5.251)
Squared Distance from upstream border to station 1 ($U1^2$)		0.186	0.181	0.152	0.190		0.409
		(0.191)	(0.180)	(0.222)	(0.191)		(0.376)
Distance follows border downstream of station 1					-0.111		
					(0.266)		
Distance follows border upstream of station 1					-0.111		
					(0.530)		
GDP of the county in which the station is located in 100000 Reis	0.354	0.341	0.249	0.221	0.341	-0.007	-0.013
	(0.345)	(0.347)	(0.341)	(0.346)	(0.347)	(0.669)	(0.661)
Population of the county, in 100,000	2.270	1.687	1.068	3.448	1.670	5.155	3.394
	(2.000)	(2.101)	(2.236)	(2.526)	(2.106)	(6.422)	(6.691)
Area of the county, in 100,000 square km	-322.743**	-274.982**	-236.013**	-259.219**	-278.598**	-319.828**	-264.123**
	(124.828)	(114.928)	(114.129)	(118.044)	(117.466)	(94.830)	(120.403)
Basin*year fixed effects	Y	Y	N	N	Y	Y	Y
Basin*month fixed effects	Y	Y	N	N	Y	Y	Y
Month* year fixed effects	N	N	Y	N	N	N	N
Basin*year*month fixed effects	N	N	N	Y	N	N	N
Station trends	N	N	N	N	N	Y	Y
Observations	5,989	5,989	5,989	5,989	5,989	5,989	5,989
R-squared	0.065	0.065	0.060	0.098	0.065	0.134	0.135
Number of pair	372	372	372	372	372	372	372
F-stat for slope of pollution function upstream = slope downstream (evaluated at 1km from border)	3.159	2.415	3.071	2.209	2.427	0.735	2.046
Prob > F	0.076	0.121	0.081	0.138	0.120	0.392	0.153
F-stat for slope of pollution function upstream = slope downstream (evaluated at 3km from border)	2.812	2.841	4.219	3.007	2.833	0.017	2.232
Prob > F	0.094	0.093	0.041	0.084	0.093	0.896	0.136
F-stat for slope of pollution function upstream = slope downstream (evaluated at 5km from border)	2.574	3.163	5.970	3.649	3.125	0.038	2.084
Prob > F	0.109	0.0761	0.015	0.057	0.078	0.845	0.150

Robust standard errors in parentheses

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

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 upstream 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 3: Station Pair Regressions: Change in Pollution from Upstream to Downstream Station Over time

Dependent Variable: Log Difference in BOD (Dowstream minus Upstream)*100

Number of borders crossed between station 1 and station 2	3.200*	3.595**	3.418**	3.886**
	(1.647)	(1.536)	(1.494)	(1.550)
Distance from station 1 to Downstream border (1D)		1.526*	1.771*	1.740*
		(0.885)	(1.029)	(1.028)
Distance from Upstream border to station 2 (U2)		-3.101	-3.347	-3.442
		(1.965)	(2.108)	(2.097)
Outside Station Pair Control Variable: Distance from Upstream border to station 1 (U1)		-1.676	-1.597	-2.179
		(2.244)	(2.244)	(2.257)
Outside Station Pair Control Variable: Distance from station 2 to its Downstream border (2D)		-0.115	0.119	0.076
		(0.522)	(0.606)	(0.609)
Distance river follows the border to upstream station			0.401	0.590*
			(0.331)	(0.344)
Distance river follows the border to downstream station			1.472	1.302
			(2.274)	(2.273)
Distance river follows the border between stations			0.403	0.320
			(0.640)	(0.624)
Distance river follows the border downstream of downstream station				-0.000
				(0.001)
Distance river follows the border upstream of upstream station				-0.001
				(0.001)
GDP of the county in which the upstream station is located in 100000 Reis	0.027	-0.233	-0.224	-0.252
	(0.693)	(0.650)	(0.674)	(0.678)
GDP of the county in which the downstream station is located in 100000 Reis	-0.048	0.009	0.001	0.006
	(0.195)	(0.206)	(0.205)	(0.209)
Population of the upstream county, in 100,000	-6.456	-4.722	-4.024	-3.973
	(7.469)	(7.162)	(7.492)	(7.509)
Population of the downstream county, in 100,000	19.427***	24.111**	20.912**	21.507**
	(5.399)	(9.619)	(9.979)	(10.276)
Area of the upstream county, in 100,000 square km	136.379	234.667	-8.623	7.237
	(210.705)	(204.781)	(438.153)	(430.683)
Area of the downstream county, in 100,000 square km	-144.256**	82.426	114.246	145.269
	(68.011)	(195.500)	(227.955)	(230.920)
Average GDP in intermediate counties in 100,000 R$ in constant 2000 R$	0.001	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)
Average Population in intermediate Counties in 100,000 people	-0.000	-0.009	-0.011	-0.012
	(0.009)	(0.014)	(0.014)	(0.014)
Average Area in 100,000 square kilometers	-0.230***	-0.190***	-0.207***	-0.180**
	(0.049)	(0.047)	(0.067)	(0.077)
River basin*year FE	Y	Y	Y	Y
River basin*month FE	Y	Y	Y	Y
Station pair Trends	Y	Y	Y	Y
Observations	5,989	5,989	5,989	5,989
R-squared	0.120	0.122	0.122	0.122
Number of pair	372	372	372	372
F-test for slope of pollution function upstream of border = slope downstream of border		3.150	3.014	3.122
Prob > F		0.0773	0.0840	0.0786

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 which there were no intermediate counties for GDP variables (cases of 0 or 1 crossings).

Table 4: Non-linearity in the Pollution Function, at Varying Distances from Borders

	Split at x=1	Split at x=3	Split at x=5	Split at x=15
U2 - far from border (Distance from upstream border to station 2 * beyond *x* km of station)	-2.523	-3.129	-2.624	-1.290
	(1.990)	(1.957)	(1.987)	(1.460)
U2 - close to border (Distance from upstream border to station 2 * within *x* km of station)	-9.510	-12.887***	-4.347	-4.040
	(18.927)	(4.621)	(4.262)	(3.509)
1D - close to border (Distance from station 1 to downstream border within *x* km of station)	9.372	7.765	0.446	-0.734
	(15.566)	(4.722)	(2.235)	(2.118)
1D - far from border (Distance from station 1 to downstream border * beyond *x* km from station)	1.941*	2.125**	1.794*	1.609*
	(1.066)	(1.000)	(1.075)	(0.956)
2D Distance from station 2 to downstream border	-0.025	0.124	-0.009	-0.916
	(0.485)	(0.501)	(0.494)	(0.799)
U1 Distance from upstream border to station 1	-2.452	-2.369	-2.453	-2.158
	(2.278)	(2.197)	(2.310)	(2.275)
Observations	5,989	5,989	5,989	5,989
R-squared	0.120	0.121	0.120	0.120
Number of pair	372	372	372	372
F-stat for slope upstream of border = slope downstream of border (U2 close = ID close)	0.628	8.418	0.808	0.558
Prob > F	0.429	0.00410	0.370	0.456
F-stat for test of non-linearity in slope (1D close = 1D far)	0.230	1.442	0.565	1.865
Prob > F	0.632	0.231	0.453	0.173

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. 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 which there were no intermediate counties for GDP variables (cases of 0 or 1 crossings).

GDP, population, area controlled in all specifications but not shown

Table 5: 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 upstream 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 6: 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).

Table 7: Causes of and Responses to Water Pollution

Panel A. County-Reported Causes of Water Pollution	Count*
Mining	235
Oil and gas from boats	81
Animal Waste	832
Materials from the Processing of Sugar	160
Industrial Dumping	521
Domestic Sewage	1595
Poor Solid Waste Management	821
Poor enforcement of river pollution regulations	648
Poor enforcement of underground water rights licensing	228
Use of Pesticides and Fertilizers	901
Others	160
Total Counties reporting Water Pollution	2121
Panel B. County Actions to Reduce Pollution	
Fining Households with Inadequate Sewer Systems	2462
Fining Companies with Inadequate Industrial Waste Management System	1007
Monitoring of Potentially Polluting Industrial Activities	596
Taxing Mining Industries	1027
Taxing Automobiles	104
Management of Toxic Waste	483
Trash Collection Program	1654
Recycling Program	1082
Creation of Sewers	1949
Other	564

*Counts are as of 2002. There were 5,560 counties in Brazil in 2002. Source: IBGE Environmental Census

Table 8: Non-point Source Pollution Measures (Placebo Check)

	Log Total Dissolved		Log Turbidity		Log Conductivity	
	Top 1% dropped		Top 1% dropped		Top 1% dropped	
Distance from station 1 to Downstream border (1D)	-0.833	0.163	-0.756	-0.246	-0.604	-0.358
	(1.725)	(1.801)	(1.400)	(1.861)	(0.928)	(1.509)
Squared Distance from station 1 to Downstream border ($1D^2$)	-0.027	-0.018	0.074***	0.055**	-0.019	-0.031
	(0.024)	(0.024)	(0.025)	(0.028)	(0.016)	(0.026)
Distance from Upstream border to station 1 (U1)	-2.486	-1.827	0.545	1.094	2.468	2.849
	(3.436)	(3.615)	(5.035)	(6.689)	(2.696)	(3.265)
Squared Distance from upstream border to station 1 ($U1^2$)	0.395*	0.284	0.036	-0.356	-0.326*	-0.164
	(0.235)	(0.238)	(0.278)	(0.424)	(0.192)	(0.179)
GDP of the county in which the upstream station is located in 100000 Reis	0.146	0.121	0.363	1.007***	0.262	-0.041
	(0.185)	(0.207)	(0.634)	(0.228)	(0.300)	(0.201)
Population of the upstream county, in 100,000	1.740	-0.017	-1.479	-14.056**	-1.452	-4.173
	(2.722)	(3.699)	(7.015)	(5.844)	(4.053)	(6.149)
Area of the upstream county, in 100,000 square km	50.663	-145.122	572.230	828.424**	-262.508	-118.067
	(194.715)	(387.520)	(427.286)	(321.383)	(293.611)	(372.832)
Includes station pair trends	N	Y	N	Y	N	Y
Observations	4,882	4,882	5,725	5,725	5,651	5,651
R-squared	0.212	0.276	0.367	0.421	0.133	0.228
Number of pairs	300	300	363	363	363	363

Robust standard errors in parentheses

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

All regressions include station pair fixed effects. Standard errors are clustered by the upstream 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).

Figure 1: Model of a River

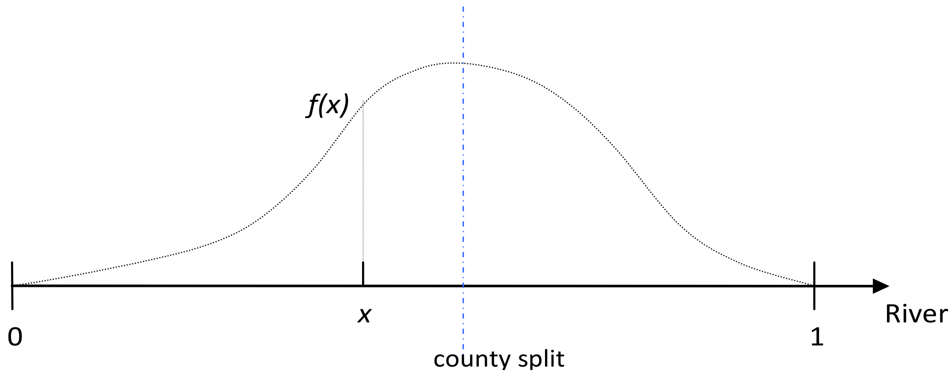


Figure 2: Emissions (q_x) Allowances Before and After the County Split

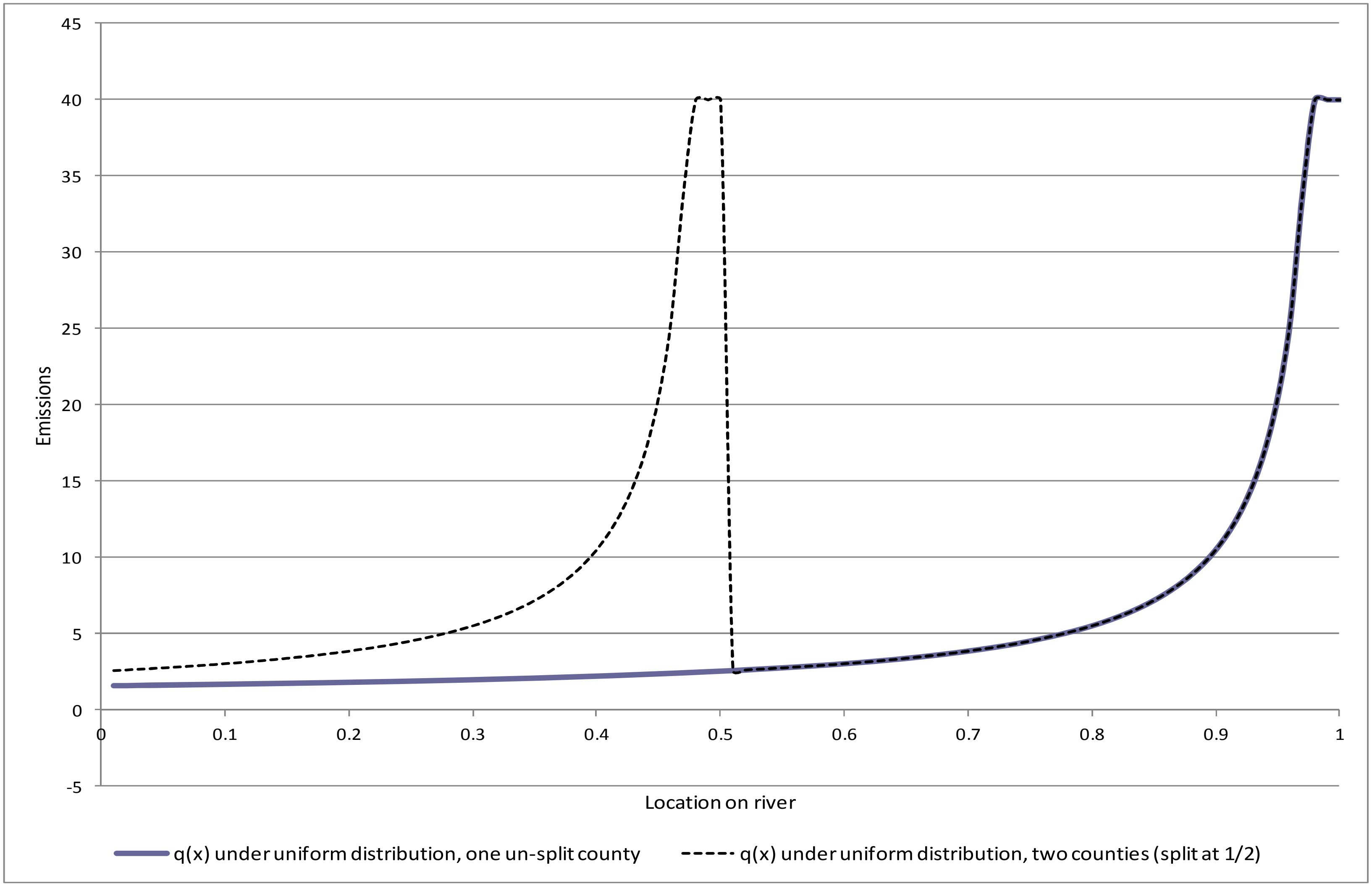


Figure 3: **The Pollution Function *P(y)* for a County with Uniformly Distributed Population before and after the Jurisdictional Split at 0.5**

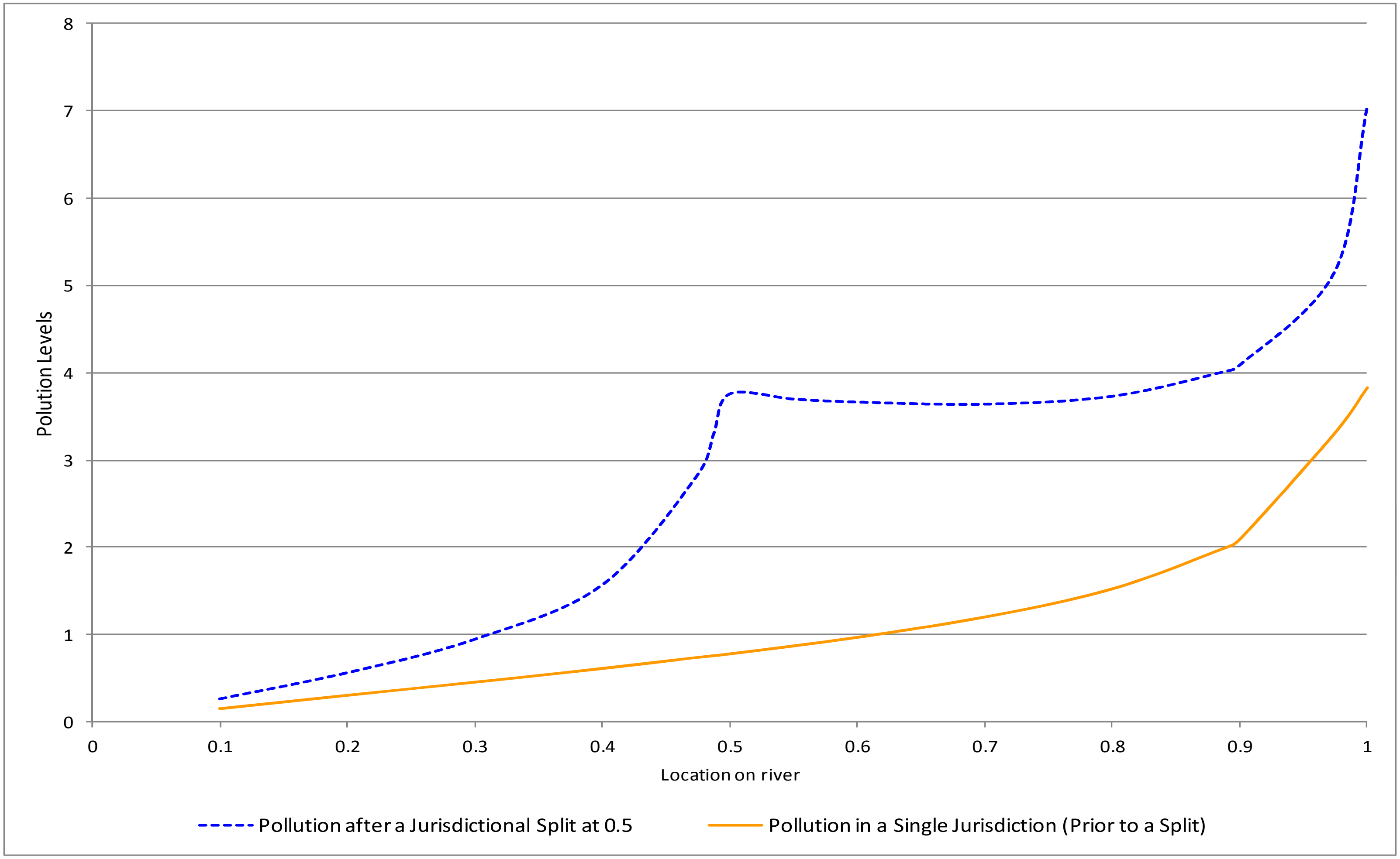
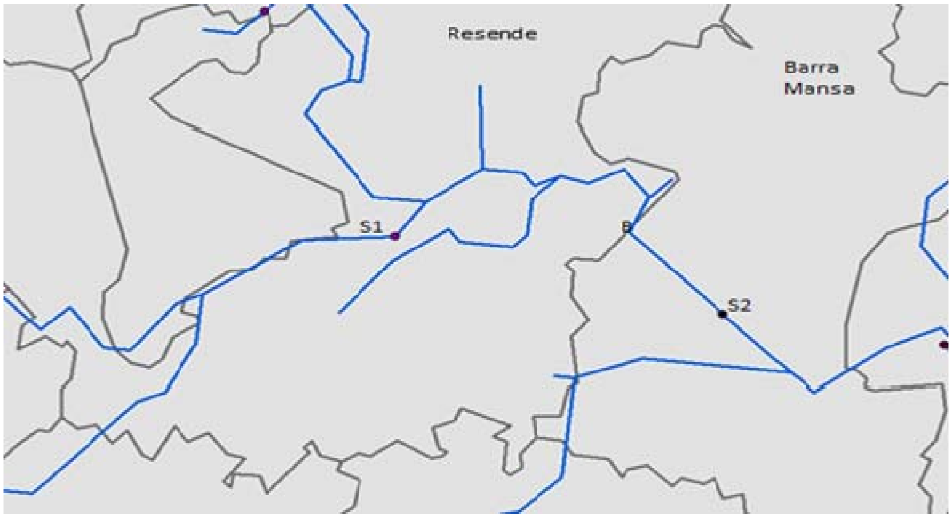
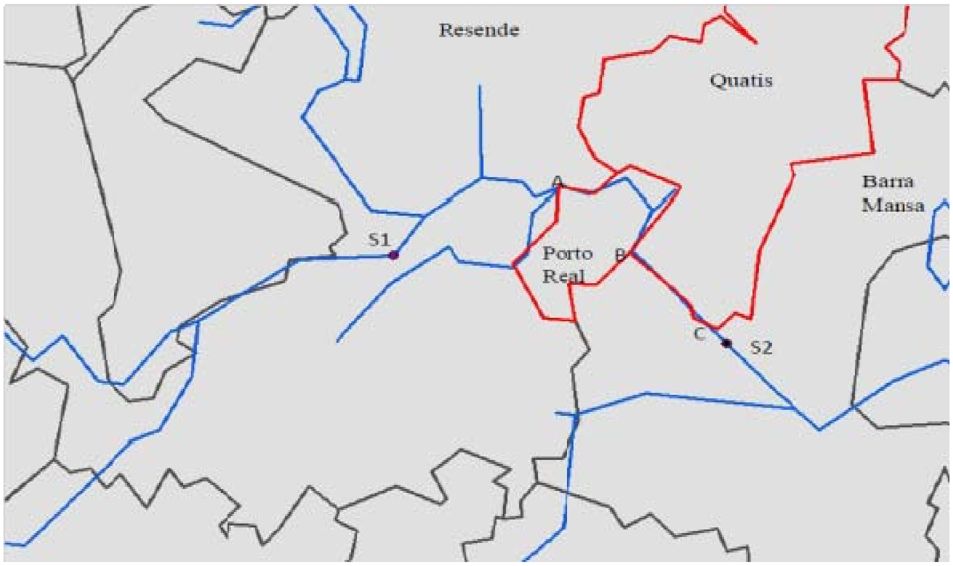


Figure 4: Example of the Evolution of County Boundaries in the State of Rio de Janeiro



1991 Map, IBGE



1997 Map, IBGE

Figure 5: Notation on Distances

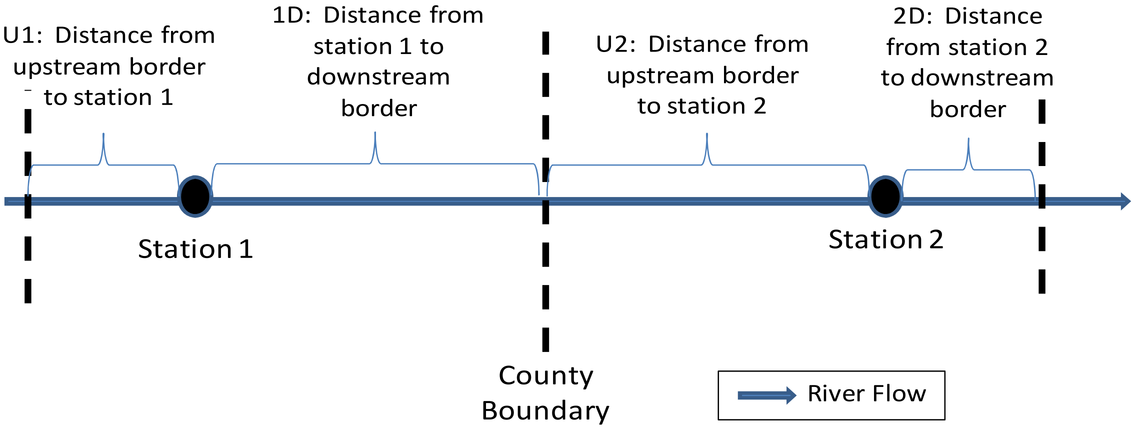


Figure 6: Rivers and Water Quality Monitoring Stations

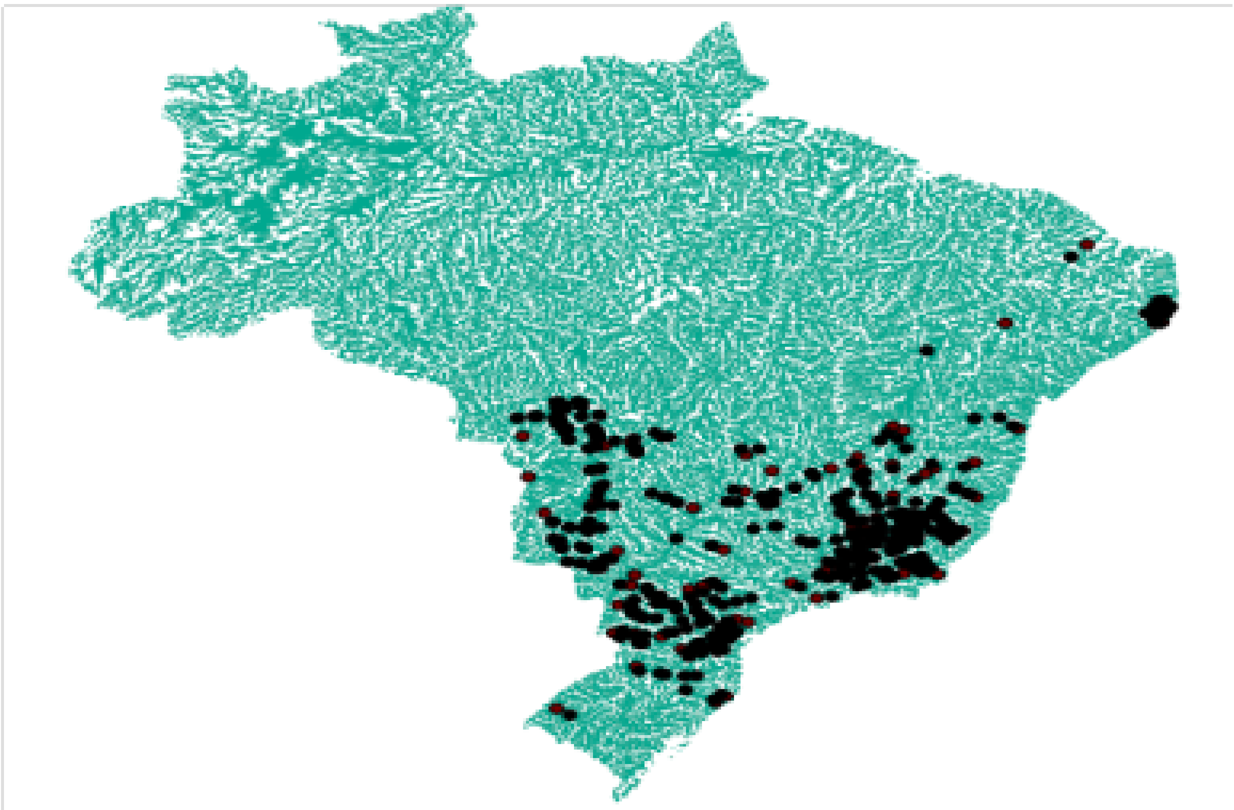
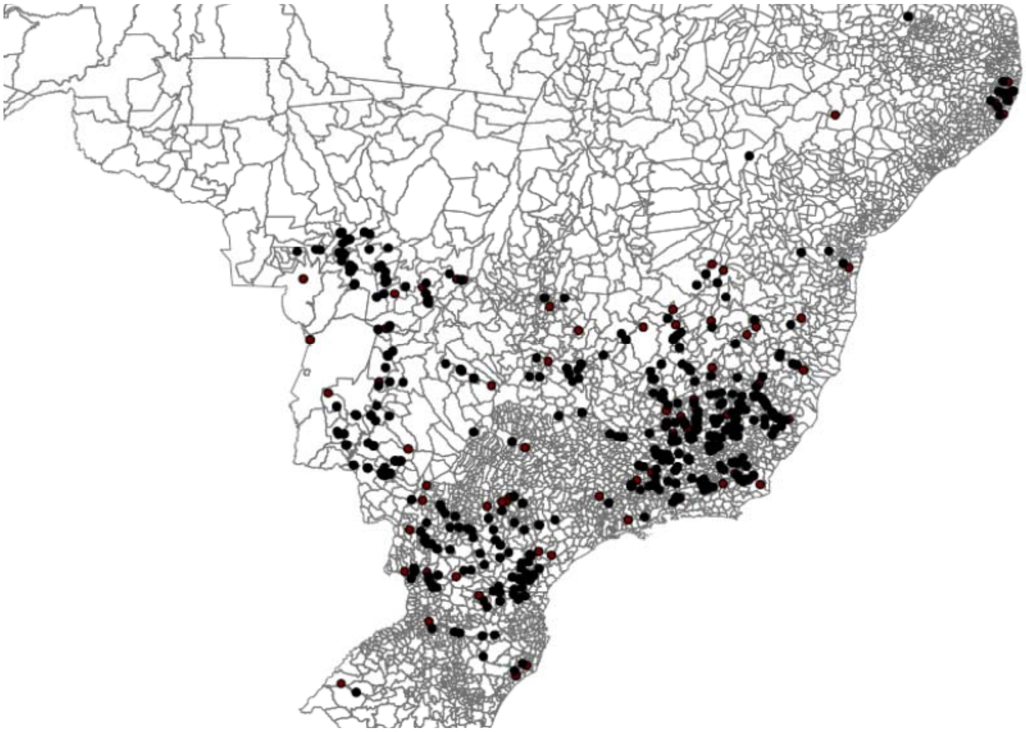


Figure 7: Water Quality Monitoring Stations and County Boundaries



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