

Statement on past, current and future research

Andrea Szabó

August 2026

My work is at the intersection of industrial organization [1, 2, 3, 4, 8, 9, 10, 13, 14] and development economics [1, 2, 4, 5, 6, 7, 8, 9, 10, 11, 12, 14]. I use methods from industrial organization, particularly demand estimation (both static [1, 3, 4, 8] and dynamic [9]), production function estimation [2, 10, 13], the estimation of static games [8] and single agent dynamic models [10, 14], and study a variety of sectors, including public utilities [1, 3, 5, 7, 11, 13], manufacturing [2, 10], retail [9], and even politics [4, 8]. Several of my papers address environmental and conservation issues [1, 5, 7, 11]. Six of my papers use original data I collected in the field. I have conducted an RCT [5], used conjoint surveys [6, 3] and fielded five surveys in Africa [1, 5, 6, 9]. I currently have ongoing projects in Ghana, South Africa, and Bangladesh. Two key themes that underlie much of my work are (i) a focus on firms, which are often neglected in the development literature, and (ii) the potential of standard economic interventions such as pricing, firm regulation, and information provision to improve welfare in developing countries.

Below I briefly describe some of my published work, and provide more details on my ongoing and planned projects.

Water related issues in developing countries

Papers [1] and [5] are the product of my cooperation with a South African water utility. In [1] I use a unique dataset of consumer billing data to study South Africa's Free Basic Water Policy, under which households receive a free water allowance. I estimate residential water demand, evaluate the welfare effects of free water, and provide optimal price schedules. I show that the free allowance acts as a lump-sum subsidy, without large effects on water consumption. At the same time, I also show that it is possible to reallocate the current subsidy to form an optimal tariff without a free allowance that would increase welfare while leaving the water provider's profit unchanged. This optimal tariff would also reduce the number of households consuming low quantities of water, a desirable policy goal according to the World Health Organization.

While working on this project, I became aware of an issue, the nonpayment of water bills by consumers, which is both crucial to many developing country providers and at the time was vastly understudied in economics. Conventional enforcement strategies often do not work. In [5], we explore an alternative strategy to reduce nonpayment: providing information. Purchasing and using water is a complicated activity: quantities used by various appliances are not directly observed by the household, and consumption and payment occur at different times. Providing information about various aspects of water consumption could improve the households' water management and consequently lower waste and their monthly bill. To study these questions, we implement and evaluate a randomized water education campaign. We estimate substantial short-run treatment effects: on the order of a 25% increase in payments over a three-month period. Interestingly, the

evidence shows that the treatment did not operate by increasing consumers' information. Instead, households may have reciprocated the provider's efforts by paying more. Our findings provide evidence that strategies other than increased enforcement can lower nonpayment.

Public utilities in developing countries around the world deliver basic services to millions of disadvantaged households, while their operation and regulation faces numerous challenges. Economic analyses of many of these questions are scarce, and I plan to continue working on this research agenda in the coming years.

Demand and supply of political candidates

In [4] and [8], we present a novel application of tools from Industrial Organization to the political market in a developing country setting, focusing on India. We argue that voters' choices between political candidates can be understood using similar methods to those used to study consumers' choices between different products.

In [4] we apply this idea to study the impact of a voting reform: the introduction of a None of the Above (NOTA) option on Indian ballots. Our findings reveal that, in the absence of NOTA, most protest voters either abstain or disperse their votes among multiple candidates, having little impact on election results. The introduction of NOTA, however, provides a clear outlet for voter dissatisfaction, which increases political participation.

To the extent that all participation in a democracy is valuable, our finding that a NOTA option can increase voter turnout is significant in its own right—particularly since this policy is relatively simple to implement.

Paper [8] builds on our model of voter demand for candidates and explores how Indian political parties select candidates, including those with criminal backgrounds. To estimate voter preferences, the study employs a discrete-choice model describing voter behavior in Indian elections. On the supply side of politics, the paper uses a game-theoretic model to understand the strategic interaction between parties as they select candidates to represent them.

The analysis reveals several key insights. While parties aim to win elections, they often prefer candidates who are not overly popular and do not threaten current party leaders' positions. Candidates with criminal histories can be selected because they are relatively easy to recruit or for strategic reasons. Simulations show that banning candidates with criminal backgrounds can alter the electoral landscape, increasing the share of educated candidates. Interestingly, both major parties could benefit from such a ban, as it would allow them to compete with more desirable candidates.

Together, these two studies offer a nuanced understanding of voter behavior and party strategies in India's electoral landscape, and illustrate that methods from Industrial Organization can usefully be employed to analyze the political market.

Research on national parks

I have a special interest in studying the economic effects of national parks, including income generated through tourism, which can support local economic development in both developed and developing countries.

The US National Park System (NPS) is the largest and best known national conservation entity in the world. Its mission is to conserve certain areas while making them accessible to the public in a sustainable manner, and this model has been adopted by over 100 countries that have opened national parks in the past 150 years. Perhaps surprisingly, no previous study has attempted to measure the causal effect of National Parks and other NPS parks on the local economy. In [7] and [11], we use a newly assembled, comprehensive dataset on all parks in the NPS, which allows us to conduct such an analysis.

In [7] we find that both the granting of “National Park” designation and the opening of a new park raise employment and income in the local economy. The magnitude of the effect is 4-6 percent, and the full increase takes 3-4 years to realize. Interestingly, these effects are not restricted to National Parks – opening parks with less well-known designations also has positive economic impacts on average.

The evidence shows that direct government spending on or around the parks cannot account for these income and employment effects. The results also cannot be explained by the diversion of economic activity from neighboring counties or from other parks, nor by changes simply in the size of the conserved areas. Instead, the evidence is consistent with economic impacts being driven by visitors.

The finding that visitation is crucial to realizing the positive economic impacts of conservation has implications beyond the NPS. For example, there are 372 national parks in 54 countries on the African continent, but many of these do not have facilities or even allow visitor access. Only a few parks appear to have visitor numbers available. Our results suggest that improving access may be important to realize the economic potential of these areas, and collecting systematic visitor data may be important to measure their impacts.

Our on-going work in [11] focuses on the heterogeneity of the economic costs and benefits of environmental conservation in the U.S. National Park System (NPS). Building on our previous paper that examined parks’ overall impact on employment and income, this project assesses the heterogeneous effects of parks on local economies. Specifically, we analyze how parks generate economic benefits through tourism in sectors like hotels, retail, and construction, while also creating restrictions that affect industries such as logging, mining, and agriculture.

Our preliminary results show that opening new parks leads to mine closures. At the same time, there is little evidence that this reduced total mining employment – perhaps because the remaining mines were able to soak up most of the workers displaced from the mines that shut down. As expected, changing an existing park’s designation to National Park has no additional negative impact on the mining sector.

We do not find negative effects on either farming or logging / forestry from either National Park designation or the opening of new parks. The evidence shows that this could be due in part to adaptation by affected firms. For example, we find that the opening of a new park leads to an increase in the volume of timber extracted in nearby areas, indicating that loggers are able to substitute towards new sources of timber when an area becomes protected.

We also estimate potential negative externalities: we do not find evidence that parks increase pollution, but there is evidence of an increase in traffic fatalities following the opening of a park, which is in line with the increase in tourism documented in [7].

Impacts of the banking system in developing countries

A different strand of my work studies the impact of banking sector expansion in developing countries.

Paper [2] studies inefficiencies in the operation of manufacturing firms in Ghana due to financial frictions. I measure firm level inefficiencies in input use among manufacturing firms by explicitly estimating their production function. I find that the fraction of undercapitalized firms is 40%, but overall, firms use 83% more capital and 40% less labor than would be optimal. Firms with formal loans and firms with more human capital are closer to their efficient capital stock. The findings suggest large potential gains in value added from adjusting input use in the optimal direction.

Paper [10] continues this research by investigating a source of potential inefficiency at the aggregate rather than the firm level. Family financing through loans for investment or intermediate input purchases may allow relatively unproductive firms to stay in the market, reducing average productivity in the economy. To quantify this effect, we estimate a dynamic model of firm behavior using data from the Ghanaian Manufacturing Survey. We supplement this data with new measures of credit market access that we assembled by collecting old phone directories. A counterfactual analysis with no family financing indicates an average productivity gain of about 10% over 20 years relative to a situation where all firms have access to family loans. This increase in productivity is accompanied by large gains in average output produced.

Because improving formal lending may reduce the availability of family loans, these results suggest a new channel through which improving credit market conditions may increase productivity in developing economies.

To confirm our simulation results in the data, in summer 2025 we located and collected information on all firms in the original sample. This information on the long-run outcomes of firms provides a unique opportunity to evaluate simulations conducted on a 30-year time horizon.

One of my on-going research projects, [14], ties to [10] by documenting the evolution of the banking system in Bangladesh and investigates the impact of several financial inclusion reforms aimed at expanding banking services in rural areas. Under these policies, the central bank mandated that for every new urban branch a bank opens, it must first open a branch in a rural area. Over the decades, the central bank implemented several variations of this regulation, adjusting the required ratio of rural to urban branches multiple times. These policies, designed to address the profitability imbalance between urban and rural branches, placed a cost burden on banks seeking urban expansion.

Using a unique panel dataset of bank branches from 2000 to 2018, we examine whether these policies succeeded in increasing rural banking coverage. Preliminary results suggest that the policy led to an increase in rural branches during the initial phase. However, once rural bank branches reached a certain level, the regulation had an unintended effect: rather than opening more rural branches, banks chose to open fewer new urban branches, resulting in an overall decline in the growth of the total number of branches.

To assess the full effect of these regulations, we have completed a major data collection effort to assemble and digitize balance sheet and other information from bank annual reports. This research will provide important insights into the effectiveness of financial inclusion policies in developing economies and their implications for the structure of the banking sector.

Research on grocery stores in developing countries

Following the theme of my research on water pricing, paper [9] investigates welfare-improving changes to firm behavior - here in the context of grocery stores' product selection.

Storable goods like laundry detergent, bathroom tissue, or non-perishable food items are often sold in the same store in widely different package sizes. The unit price for these goods is typically lower if the consumer buys a larger package. While buying small packages can have some benefits, if smaller packages have a larger unit price, then households overspend on these storable goods. In [9], I use a unique scanner dataset on laundry detergents from Unilever South Africa to estimate a dynamic demand model that allows quantifying the determinants of these choices.

I show that the combination of non-linear pricing and temporary promotions can result in situations where consumers are better off buying two smaller packages instead of a larger one. Ignoring the presence of these “bundling opportunities” can lead to very misleading conclusions regarding consumers' preferences for small package sizes in developing countries. I estimate that many of the households buying the smallest package sizes are in fact creating bundles and purchasing a larger total quantity. On the median market, there are over 1.5-3 times as many households who purchase 5 kg of detergent as households who purchase a single 5 kg package.

The model also provides a measure of the impact of different package sizes on consumer welfare. This is relevant in the African context, where grocery store chains are expanding in low-income neighborhoods by opening small stores that do not carry all available package sizes. I find that when small stores are limited in the selection of package sizes they can carry, larger package sizes tend to create higher consumer welfare than small package sizes.

On-going research on public utilities

Project [13] examines the efficiency and productivity of postal services, with a particular focus on how these services are provided when profit maximization is not the primary objective. Governments offering essential services like electricity or postal systems often set prices administratively, which can lead to a disconnect between prices and the actual costs incurred by the firms, especially when prices remain fixed for extended periods. In the postal sector, the aim is to ensure affordable, widespread access to high-quality services, as seen in the EU Postal Services Policy.

Our research addresses the challenges of applying traditional production function estimation methods, which are commonly used for profit-maximizing firms in the manufacturing sector, to the postal sector. Postal firms produce multiple outputs—letters and packages—using the same set of inputs. In this context, inputs can be complements in production, even though the products themselves are not substitutes for consumers. In addition, inputs are shared across multiple products. By extending the standard methods to incorporate these elements, this research aims to provide a more accurate assessment of efficiency and productivity in postal services in different countries, offering insights into the broader implications of non-profit-maximizing industries.

On-going data projects

I am currently working on two major data collection projects. I will use these data projects in multiple future studies, and they will also provide a valuable resource for other researchers interested

in exploring different aspects of the data.

Banking System in Bangladesh. Bangladesh allowed the establishment of its first private bank in 1983. I am leading a data collection effort to gather annual reports from all private banks in the country from their inception to the present. While these reports are technically public information, they are not readily accessible, and no institution, including the Central Bank of Bangladesh, holds a complete collection of these documents.

We initially attempted various methods to collect electronic copies, such as downloading them from the internet or purchasing them from the Dhaka Stock Exchange. We also approached individual banks directly, invoking the Public Information Act to collect their annual reports. In addition, I hired several students from Dhaka University to scan available copies from various libraries, such as the Library of Bangladesh Bank and the Bangladesh Institute of Banking Management. As of now, we have acquired all of the approximately 1,000 annual reports of the country's private banks.

We are now compiling a database with key variables, such as the number of bank branches, employment figures, total operating income and expenditure, total assets, deposits, and other financial indicators available in the reports. Throughout this process, we are documenting accounting changes and any errors in the reports to provide detailed context for future users of the data. To my knowledge, there is no other dataset that offers comprehensive information on an entire country's banking system from its inception. I plan to use this data for several projects, one of which I describe above [14].

National Park System (NPS) Data Collection. Building on our earlier work, I am working on a data collection project on the National Park System (NPS), which will provide further insights into the economic and environmental impacts of U.S. national parks. In order to better understand the heterogeneity in these impacts, we will extend our data collection efforts to include more granular information on park amenities, the road networks leading to park entrances, and city-level information on different industries. This project includes gathering historical data from sources such as AAA tourbooks and state timber auction records. As with the banking system project, this data will be made publicly available for future research, supporting a wide range of studies on conservation, economic development, and environmental policy.

Publications

- [1] Andrea Szabó. "The value of free water: analyzing South Africa's free basic water policy". *Econometrica* 83.5 (2015), pp. 1913–1961.
- [2] Andrea Szabó. "Measuring Firm-Level Inefficiencies in the Ghanaian Manufacturing Sector". *Economic Development and Cultural Change* 66.3 (2018), pp. 447–487.
- [3] Andrea Szabó and Vinh Pham. "Net neutrality and consumer demand in the video on-demand market". *Information Economics and Policy* 61 (2022), p. 100993.
- [4] Gergely Ujhelyi, Somdeep Chatterjee, and Andrea Szabó. "None of the Above: Protest Voting in the World's Largest Democracy". *Journal of the European Economic Association* 19.3 (2021), pp. 1936–1979.
- [5] Andrea Szabó and Gergely Ujhelyi. "Reducing nonpayment for public utilities: Experimental evidence from South Africa". *Journal of Development Economics* 117 (2015), pp. 20–31.

- [6] Andrea Szabó and Gergely Ujhelyi. “Choice and happiness in South Africa”. *Economics Letters* 155 (2017), pp. 28–30.
- [7] Andrea Szabó and Gergely Ujhelyi. “National parks and economic development”. *Journal of Public Economics* 232 (2024), p. 105073.

Working papers

- [8] Arvind Magesan, Andrea Szabó, and Gergely Ujhelyi. “Do Parties and Voters Disagree? An Equilibrium Analysis of Candidate Selection in India”. *conditionally accepted at Econometrica* (2026). URL: https://www.uh.edu/~aszabo2/AAG_web.pdf.
- [9] Andrea Szabó. “Bundling to save: Analyzing package size choices in South African grocery stores”. *Working paper* (2026). URL: https://www.uh.edu/~aszabo2/szabo_paper.pdf.
- [10] Andrea Szabó and Gergely Ujhelyi. “Family Financing and Aggregate Manufacturing Productivity in Ghana”. *Working paper* (2015). URL: <https://uh.edu/~aszabo2/ghana16.pdf>.
- [11] Andrea Szabó and Gergely Ujhelyi. “Evaluating Industry Impacts and Externalities of the US National Park System”. *Working paper* (2023). URL: <https://uh.edu/~aszabo2/npsB.pdf>.
- [12] Tolga Umut Kuzubas and Andrea Szabó. “Job search through weak and strong ties: Theory and evidence from Indonesia”. *Working paper* (2017). URL: https://uh.edu/~aszabo2/paper_2017.pdf.

On-going projects

- [13] Andrea Szabó and Dietrich Vollrath. “Beyond Profit: Measuring Efficiency and Productivity in the Postal Sector”.
- [14] Md Shahadath Hossain, Nabila Rahman Biju, and Andrea Szabó. “Banking Expansion in Bangladesh: The Impact of Rural-Urban Branch Regulations”.