Statement of Research and Teaching • Daniel Björkegren • October 2022

Research

I work on the provision of digital technologies, with a focus on developing economies. These technologies raise new policy questions, enable new measurement, and enable new modes of action (through automated decisionmaking). I work to connect the loop between measurement and action, by combining theory and computational techniques.

While my work is inspired by the lives of the world's poor, I try to find principles that are relevant everywhere. I develop partnerships with private firms and governments to obtain rich, digital data on populations that are sparsely documented in survey and traditional administrative data. When suitable, I supplement this with custom surveys, or run randomized experiments. I contribute to the intersections of development economics, the industrial organization of digital technology, and applied machine learning.

> A first part of my work analyzes the industrial organization of networks using digital data. Mobile phones have spread dramatically, knitting even the poor into the global economy. My work analyzes the value mobile phone networks provide, what drives consumers to adopt, and how the industry should be regulated. I also consider another important network industry in emerging cities: informal transit.

> A second part of my work is on applied machine learning. The spread of computation has enabled decisionmaking to be automated. My work has provided foundations for one early such application: digital credit scoring for people excluded from traditional banking. I also explore how such algorithms can be made more humane: robust, transparent, and better aligned with societal values.

> A third part of my work explores how digital data can measure human behavior in detail new to the social sciences.

1. Industrial Organization of Networks

1a. New Mobile Phone Networks and Direct Network Effects

Mobile phone networks increasingly act as gatekeepers to information and finance across developing economies. Yet, there is little work on adoption, the supply of service, or regulation. How should societies tax networks, ensure coverage to remote areas, and ensure sufficient adoption? Should societies tolerate monopoly networks, or split them up?

A primary challenge in answering these policy questions is that phone networks produce direct network effects. As more of the people you wish to interact with adopt phones, it becomes more beneficial for you to adopt. Because of these linkages, a policy that affects one person's adoption can cause ripple effects through the entire network. These effects are similar to those that characterize other important tech platforms (such as social networks, messaging platforms, and payment systems). Because developing country networks are still growing, these effects are more important than in mature phone networks.

These ripple effects make it difficult to analyze policy in any direct network industry. First, it is difficult to gather data on the entire network over which effects may spread. Second, even with this data, it is difficult to learn how one person's decision will affect another's. Two individuals who are connected may adopt at similar times because they value interacting with each other, or because connected people are exposed to similar shocks. Because phones are durable it matters not only which contacts have a phone at a given point in time but also who is likely to purchase a phone in the future, which may depend in turn on who whether you purchase a phone in the first place. There thus tend to be multiple equilibria. There is an empirical literature analyzing policy in simpler *in*direct network industries, in which consumers benefit from additional users not because they value links with them, but because popular platforms are better served by the other side of the market (such as two sided markets, where a rider cares about how many Uber drivers there are). However, there is very little work using data to evaluate policy in industries with direct network effects.

The article, **The Adoption of Network Goods: Evidence from the Spread of Mobile Phones in Rwanda** (*Review of Economic Studies*, 2019), overcomes these challenges to estimate and simulate the adoption of nearly every mobile phone subscriber in Rwanda from 2005-2009, using 5.3 billion transaction records. Because phone calls were billed by the second, a subscriber must value each connection at least as much as the cost of calls placed across it. Further, because the firm changed calling prices and increased the quality of service, I can identify the underlying demand curve for communication across each link—and thus the value of each link in the network. I develop a simulation method that allows each consumer to adjust their adoption and usage in response to a policy and to each other, capturing how the effects of a policy ripple through the network and across physical space. I characterize multiple equilibria in adoption using supermodularity. This first article uses this method to produce three empirical results:

First, it provides the first micro-identified estimates of the net welfare generated by a developing country mobile phone network: the equivalent of 2-3% of Rwanda's GDP over this period, with 51% accruing to consumers, 14% to government, and 35% to operators.

Second, it finds that taxing the adoption and use of growing digital services can be extremely distortionary. The mobile industry contributed an average of 7% of government revenues in sub-Saharan Africa (GSMA, 2008). Standard measures used by industry would suggest that these taxes induce a welfare cost to Rwanda of $1.22 per dollar of government revenue, but accounting for network effects, welfare costs are 2.4 to 4.5 times larger.

Third, it finds large spillovers from serving rural areas. When the operator was required to build towers in remote regions, 78% of the consumer benefits accrued to individuals in other regions. This suggests that policies that rely on communities to fund their own coverage are unlikely to adequately cover remote areas. Regulations that required the operator to provide service to remote areas boosted welfare.

The model in this article also provides the foundation for two additional articles.

Attaining sufficient adoption in networks is a challenging problem. The article, **Network Adoption Subsidies: A Digital Evaluation of a Rural Mobile Phone Program in Rwanda**, (with Burak Ceyhun Karaca, *Journal of Development Economics*, 2022) analyses the effects of a government adoption subsidy. Most subsidized handsets remain in rural areas where the subsidy was targeted, and subsidized handsets are used as much as handsets purchased at retail prices. Simulations suggest the enacted handout had a positive social impact, but that an alternate policy of providing vouchers may have had a higher impact.

A second problem in network industries is that they in many cases tend towards monopoly, which has led to public concern about dominant tech firms. How should societies discipline these industries? A single network may internalize more network benefits and thus invest more, but also may take advantage of consumers. It is challenging to empirically analyze competing networks. In markets that are already competitive, it is difficult to obtain and link network data from all competitors. And theoretically, each firm must choose a path of actions over time anticipating what each other will choose, anticipating what consumers will demand, where each consumer values the usage and adoption of potentially millions of other consumers.

The article, **Competition in Network Industries: Evidence from the Rwandan Mobile Phone Network** (*RAND Journal of Economics*, 2022) is to my knowledge the first analysis of competition between direct network goods that uses data on individuals. Building on the previous model, it evaluates what would have happened if the monopoly network were split in two earlier than it was, combining data on nearly the entire network from the original firm, and data on the eventual competitor from a consumer survey I fielded. It allows two competing firms to decide prices and investments in coverage, and consumers to select and switch operators. It characterizes the space of multiple equilibria by exploiting supermodularity in two dimensions. It finds that adding an additional competitor earlier could have reduced prices 30-50%, and *increased* incentives to invest in rural towers. Such a policy would have increased total social welfare by the equivalent of 1% of GDP, an enormous impact. However, incentives to invest would be lowered if competition were pushed too strongly, by forcing firms to interoperate their networks without charge (a policy proposed in the U.S.).

This body of work provides a proof of concept for digital policy evaluation. The field of industrial organization typically relies on public or survey data, which is sparse in most contexts, and nonexistent in many developing countries. These articles illustrate how digital regulators could evaluate policy given access to data that firms collect automatically as a side effect of operation. My article, **To Regulate Network-Based Platforms, Look at Their Data** (with Chiara Farronato, *Harvard Business Review* website, 2021, not peer reviewed) distills policy lessons from the above work, and makes a call for governments to guide policy for emerging tech antitrust questions using data from platforms themselves.

These articles additionally show the importance of modeling the nuanced structure of networks in the presence of direct network effects. I find that revenue estimates can be biased ranging from 52% too small to 86% too large in demand systems that do not model the full structure of the network.

Evidence of impact

Together this work shows the importance of the supply side in providing services to the poor. I have been invited to present these articles at the World Bank multiple times, and was a panelist on data infrastructure policy for the introduction of the *World Development Report* (2021). This work has influenced several multimillion dollar Gates Foundation initiatives oriented towards increasing supply of digital services, including *Financial Inclusion Through Interoperability* (Toulouse School of Economics) and *Retail Finance Distribution* (University of Ghana). I serve as an affiliate and on the scientific committee of these initiatives, respectively.

As one of few empirical analyses of policy accounting for direct network effects in any context, this work has also had an impact in the developed societies. I have been invited to present the papers on adoption and competition at the US Federal Trade Commission (FTC) and Federal Communications Commission (FCC), and they have been cited extensively in private legal disputes over network firms.

1b. Informal Transit Networks

Transit is another important network industry. In many emerging cities, transit is private: operated by many small bus companies, some so small they only own a single bus. These systems are flexible, but can be disorganized and dangerous. In recent years, many cities have invested in formal public transit, but these systems are costly and unlikely to displace informal transit anytime soon. A series of projects seek to informal transit in Lagos, which has a population of 22 million and is projected to become the world's largest city. 75% of motorized trips in Lagos are via informal transit. We are measuring the network with novel observational data, mobility measures from mobile phones, and a series of randomized controlled experiments to assess demand.

The project, **Efficiency of Informal Transit Networks** (with Nick Tsivanidis, Alice Duhaut, and Geetika Nagpal, in progress) assesses the degree to which competing informal firms provide efficient service with a structural model. We will assess whether competing transit providers overenter central routes, or provide too little service on the fringes.

The related project, **Does Public Transit Crowd Out Informal Transit** (with Nick Tsivanidis, Alice Duhaut, and Geetika Nagpal, in progress) studies how informal transit systems respond when the city introduces 820 new formal, government-owned and regulated buses across 50 routes. The study will assess how the introduction of this new travel option affects riders and operators, in collaboration with the Lagos Metropolitan Area Transport Authority (LAMATA).

2. Applied Machine Learning

I work on new applications of machine learning, and making algorithmic decisions more robust and better aligned with social preferences.

2a. Digital Credit

Mobile phones can offer digital services at close to zero marginal cost. This makes it viable to serve some of the 1.7 billion people who lack access to traditional financial services.

Because many of these people lack formal financial histories, they lack credit scores. My paper, **Behavior Revealed in Mobile Phone Usage Predicts Credit Repayment** (with Darrell Grissen; proposal posted online 2010; working paper 2015; *World Bank Economic Review*, 2020) shows that nuances captured in the use of mobile phones themselves can be used to predict repayment for unbanked borrowers. It combines phone transactions, data on repayment of credit, and any records present in the credit bureau from a Latin American country. It finds that among people with thin or nonexistent credit files, models trained on digital traces achieve comparable performance to credit bureau models.

A companion paper, **The Potential of Digital Credit to Bank the Poor** (with Darrell Grissen; *American Economic Association Papers and Proceedings*, 2018, not peer reviewed) shows that digital platforms can theoretically expand access to credit, by lowering the fixed cost of serving consumers. In simulations it finds that performance of digital credit scoring would decline only modestly for individuals with sparse digital footprints.

Digital credit products using similar forms of credit scoring have seen dramatic uptake in developing countries. This has led to widespread concern that borrowers are being taken advantage of (Donovan and Park, 2019). The paper, **Instant Loans Can Lift Subjective Well-Being: A Randomized Evaluation of Digital Credit in Nigeria** (with Joshua Blumenstock, Omowunmi Folajimi-Senjobi, Jacqueline Mauro, and Suraj Nair; working paper) provides the first experimental estimates of the welfare impact of digital credit in a developing country. It joins two quasiexperimental evaluations of digital credit in the literature (Bharadwaj et al., 2019; Brailovskaya et al., 2021). We ran a field experiment that randomized the approval and amount of thousands of loan offers in Nigeria, and later surveyed respondents to measure effects on welfare. Randomly being approved for credit substantially increases subjective wellbeing, and we rule out large positive or negative effects on financial health, resilience, and women's economic empowerment. We find little evidence of behavioral traps: if anything, borrowers overestimate future borrowing on average.

While traditional impact evaluations require surveying, with digital technologies it may be possible to measure welfare outcomes using digital trace data. Such data is already available at scale, without cost, and could automatically be fed into decisionmaking algorithms. **Nonparametric Causal Estimators for Multivariate Missing Data: An Application to Estimate Treatment Effects from Digital Trace Data** (with Jacqueline Mauro, and Joshua Blumenstock; rejected with an invitation to resubmit from *Annals of Applied Statistics*) evaluates the welfare impact of digital credit using a second randomized controlled trial, in Kenya. We find that access to short term high interest loans appears to increase savings, decrease reliance on others to pay for phone calls, but also leads to increased borrowing from other sources.

2b. Robustness and Manipulation

Manipulation Proof Machine Learning (with Joshua Blumenstock and Samsun Knight) acknowledges that machine learning models designed to describe the world can be poorly suited to making decisions. Once an algorithm is implemented, it can be gamed, which may result in decisions that are arbitrarily poor or unsafe. This can undermine the use of machine learning in critical applications, and is used to justify keeping algorithms opaque. This problem arises because the standard estimators used to construct decision rules assume that the relationship between the outcome of interest and human behaviors is stable. But this assumption tends to be violated as soon as a decision rule is implemented, which generates incentives for agents to change their behavior (Lucas, 1976). In absence of a framework, tech firms and implementers apply ad hoc solutions, like repeatedly applying estimators that assume that behavior will be stable—which are perpetually mistaken. Building on a recent theoretical literature in computer science (Bruckner and Scheffer, 2011; Hardt et al., 2016), we embed a model of behavior within a common estimator, to create an estimable method that anticipate gaming. Our approach generates rules designed to remain robust when implemented and made transparent.

Typical datasets used in supervised machine learning observe behavior in a single environment, prior to the implementation of any decision rule, or with a single static rule. Manipulation is not detectable in these datasets, because implementing a decision rule changes the environment. We introduce the term 'counterfactual fit', to denote when we care about loss in a counterfactual environment, such as where a particular model is implemented and made transparent. To estimate counterfactual fit, we develop a new platform for implementing different decision rules experimentally.

We created a new app to mimic digital credit, together with the Busara Center, and ran a field experiment in Nairobi, Kenya with 1,557 people. The app gathers the same data as a digital credit app, and experimentally varies the decision rule each individual faces each week, and whether they see the exact decision rule (transparent treatment) or just a description of what is being predicted (opaque). We use this setup to experimentally estimate the costs of manipulating different behaviors, for different types of individuals. We find that behaviors that would appear similarly predictive to a standard machine learning method can be wildly differentially manipulable, and thus would attain substantially different predictive performance when implemented.

We document that even new users of technology manipulate their digital behavior, and find that our method is more accurate than standard methods when implemented. Our framework also suggests that the sharp nonlinearities generated by many sophisticated machine learning methods are unlikely to be optimal in the presence of manipulation.

The paper also contributes to the policy debate on algorithmic transparency: a concern raised by tech firms is that if decision rules (like Google search, or Instagram's ranking) were disclosed, they would become so manipulated that they would cease to function. Our method can simulate the equilibrium change in performance if the details of a decision rule were disclosed. We find these costs are small in our context.

It also connects to a debate on the role of theory in machine learning. The field of machine learning has made substantial progress in recent years by combining atheoretical approaches with massive amounts of data (for example, deep learning). However, in applications where counterfactual fit is the appropriate measure of performance, one will rarely have sufficient data from the exact environment of interest. We show that modeling the theoretical relationship between the model and the resulting incentives to manipulate can improve performance. We thus demonstrate a form of *structural machine learning* that combines data with a theory of behavior.

2c. <u>Welfare Sensitive Machine Learning</u>

A related set of problems arise because algorithms encode explicit values. Societies are grappling with the question: what values should be encoded?

One concern is that algorithms that are tuned for one objective can cause problematic side effects. **Balancing Competing Objectives with Noisy Data: Score-Based Classifiers for Welfare-Aware Machine Learning** (with Esther Rolf, Max Simchowitz, Sarah Dean, Lydia Liu, Moritz Hardt, and Joshua Blumenstock; *International Conference on Machine Learning (ICML)*, 2020) develops a theoretical framework for optimizing multiple objectives. We show that if a primary objective (e.g., welfare) can be measured only noisily, it may be better served by optimizing a secondary objective that is better measured if it is sufficiently correlated (e.g., engagement). We apply the framework to YouTube's recommendation system. The conference article is published in a top machine learning venue; the workshop version of the paper earned a best paper award at the *Neural Information Processing Systems (NeurIPS) Joint Workshop on AI for Social Good*, 2019.

While there is substantial work across fields on how to make decisions when an exact objective function is provided, it is difficult for societies to weigh different harms and benefits in order to provide such a function. **(Machine) Learning What Policies Value** (with Joshua Blumenstock and Samsun Knight; revision requested, *Review of Economic Studies*; computer science article published nonarchival in ACM EAAMO, 2021) develops a method to infer the preferences over households and outcomes that are consistent with a policy proposal. It does so by comparing who benefits most from an intervention (using new machine learning estimates of heterogeneous treatment effects), with who it is allocated to (based on proposed eligibility criteria). We apply it to Mexico's PROGRESA anti-poverty program. While certain subgroups—such as indigenous households—were prioritized in the allocation, those groups benefited so much more from receiving the program that the policy implicitly assigned them a *lower* welfare weight. We find that the implied value of health and schooling impacts are estimated imprecisely but are consistent with conventional valuations. We additionally find that the implemented policy was by and large consistent with preferences elicited from Mexican constituents through an online survey. Our method allows societies to invert policy discussions, to debate ends (welfare and objective weights) rather than their means (eligibility criteria).

Welfare Sensitive Credit Scoring (with Joshua Blumenstock; in progress) will apply this framework using experimental estimates of heterogeneity in impacts on profit and welfare. We will illustrate how a socially conscious lender could adjust lending on the margin to improve welfare.

2c. <u>Causal Machine Learning</u>

Machine learning methods also provide new opportunities to address traditional econometric challenges. **Causal Inference from Hypothetical Evaluations** (with B. Douglas Bernheim, Jeffrey Naecker, and Michael Pollmann; submitted) develops a new method for inferring the causal effect of treatments on choices. We undo biases in hypothetical choices, by machine learning how they relate to real choices. It can recover the effects of a treatment on choices even if the treatment is assigned endogenously and standard estimation methods are poorly suited, or even if the treatment is a proposal that has yet to be implemented. It can also recover heterogeneous treatment effects more comprehensively than standard methods. We demonstrate the approach to estimate the price responsiveness of the demand for snack foods in the laboratory, and the response of contributions to the availability of matching funds on a microfinance website. The approach yields estimates close to ground truth effects recovered in experiments we ran in both settings.

Evidence of impact

My proposal for digital credit scoring was one of the first public documents proposing credit scoring based on digital traces. The paper itself was given a 4 minute segment on *NPR Morning Edition* in 2015. It has been cited in Susan Athey's piece *The Impact of Machine Learning on Economics* (Athey, 2018), and the book *The Age of Surveillance Capitalism* (Zuboff, 2019), which was rated as a top nonfiction book by the *New Yorker*.

This approach is now commonly used for applicants who lack financial histories in a new industry providing digital loans to the poor across the developing world. Firms using similar approaches have raised nearly $0.5 billion in venture capital funding, and over 68 products are available worldwide. The expansion of digital credit has spurred the Gates Foundation to fund several multimillion dollar initiatives (the Digital Credit Observatory at UC Berkeley, and the Consumer Protection Initiative at IPA).

The academic relevance of this work has been borne out by the more recent explosion in empirical work on machine learning prediction problems. I organized the session, *Economic Applications of Machine Learning*, at the American Economic Association 2018 meetings, and presented my work on digital credit scoring, accompanied by work by Susan Athey, Josh Blumenstock, and Ed Glaeser. Despite disrupted attendance to the conference due to a blizzard, the session was standing room only; people had to be turned away.

3. Big Data for Development

A common strand through my work is the use of digital data in developing economies.

During the COVID-19 pandemic, many researchers and policymakers relied on measures of mobility from smartphones to predict the spread of disease and assess compliance with recommendations. However, in many poor countries few people have smartphones. **Assessing Bias in Smartphone Mobility Estimates in Low Income Countries** (with Sveta Milusheva and Leonardo Viotti; *ACM Conference on Computing and Sustainable Societies (COMPASS)*, 2021) uses cell tower data from an operator in Uganda to track both the 16% of mobile phones that are smart and the 84% that are not. In response to a COVID-19 lockdown, smartphones reduce mobility 40-60% more than non-smartphones, suggesting common smartphone measures can yield biased measures of population movement in settings with low adoption.

Measuring Informal Work with Digital Traces: Mobile Payphone Operators in Rwanda (*International Conference on Information and Communication Technologies and*

Development (ICTD), 2020) creates a dataset of the minute by minute business decisions of informal vendors, and uses this to trace their learning curves as they enter the industry.

Many important economic processes are affected by how individuals learn the effects of different actions (including technology adoption, management practices, and everyday choices). However, it is difficult to analyze learning in real world environments. Analysts seldom have data on the history of actions and outcomes, and even when they do, seldom know what the results of actions not taken would have been. In the project, **How Do Individuals Learn: Evidence from Rwanda** (with Aislinn Bohren, Ashesh Rambachan, and Patrick Vu; in progress), we analyze how hundreds of thousands of Rwandans experimented with a new mobile phone plan that was cheaper for short calls, and was ultimately cheaper for over 85% of people. We use rich data on every learning experience in the country—including each decision to switch plans, each transaction, and each request for feedback on charges (balance inquiry). Because we can bill each transaction under either plan, we can bound the effects of actions not taken. We will use this data to test which theoretical models are consistent with real world learning.

Overall evidence of impact of research:

As overall measures of the impact of my research, I have given numerous invited seminars at institutions such as Harvard, MIT, Stanford, Princeton, Columbia, NYU, Yale, Dartmouth, UC Berkeley, the University of Chicago, and Washington University in St. Louis. These seminars have been in a variety of groups at economic departments (development, industrial organization, econometrics, finance, applied microeconomics), business schools, and computer science departments. I have presented at top conferences, including the *National Bureau of Economic Research* (NBER) 9 times, and the *Bureau for Research and Economic Analysis of Development* (BREAD) 1 time.

I regularly speak at the World Bank and have been asked to contribute to the World Development Report and to join the advisory committee for the flagship report on the digital economy for Africa.

With junior colleagues I have raised 12 federal and private foundation grants worth a combined total of over $1.3 million. Including grants with senior colleagues Michael Kremer and Shawn Cole, I have raised over $4.8 million. I have been invited to give keynote speeches at research conferences.

Through 2020 I served on the board of the nonprofit I cofounded with Michael Kremer and Shawn Cole, Precision Agriculture for Development, which currently serves over 4 million farmers with digital advisory services.

Teaching

A popular view suggests that theory is becoming less relevant: with modern datasets we can just 'let the data speak.' I developed and taught the course **Big Data** (Econ 1660), which begins with this premise and quickly reveals that it is difficult to learn from data without imposing any structure. The arc of the course then follows the question: 'what do you need to add to data in order to learn from it?' I created a series of assignments that mimic real world analyses to explore this, contrasting machine learning approaches (trees, random forests) with more parametric approaches (OLS, LASSO), addressing overfitting, uncovering the conditions under which we can interpret estimates as causal, and finally exploring strategic interaction with structural models. We explore each topic through datasets designed to mimic real decisions made by policymakers and managers. The highlight of the course is a week-long pricing competition in which students set daily prices to compete with another group, learning from historic data, economic models, and from each others' behavior. Guest speakers have included a CEO of a consumer tech startup with over 3 million users, the technical director for AI at the US Department of Defense, and a data scientist from the Baltimore Orioles. I created the course in 2016, and it was one of the first courses combining machine learning and economics at any institution. It has been featured in multiple Brown University external communications, and the course assignments posted online have influenced the design of courses at other institutions.

I additionally have taught a course on **Applied Causal Inference** (Econ 1629), developing inquiry-based assignments to illustrate core principles.

For researchers, tapping these new opportunities requires both new questions and new partnerships with the organizations generating the data. I developed and taught a semester graduate course **Development Economics** (Econ 2520) which exposes students to traditional topics using a sequence of intuitive models, and also includes unique sessions, on passively collected 'big data' and creating research partnerships. These sessions have shaped the research of several graduate students.

References

My papers (with coauthors) cited in this statement

Bernheim, B.D., Björkegren, D., Naecker, J., Pollmann, M., 2021. Causal Inference from Hypothetical Evaluations (Working Paper No. 29616), Working Paper Series. National Bureau of Economic Research.

Björkegren, D., 2022. Competition in network industries: Evidence from the Rwandan mobile phone network. The RAND Journal of Economics 53, 200–225.

Björkegren, D., 2020. Measuring Informal Work with Digital Traces: Mobile Payphone Operators in Rwanda, in: Proceedings of the 2020 International Conference on Information and Communication Technologies and Development, ICTD2020. Association for Computing Machinery, pp. 1–5.

Björkegren, D., 2019. The Adoption of Network Goods: Evidence from the Spread of Mobile Phones in Rwanda. Review of Economic Studies 86, 1033–1060.

Björkegren, D., 2010. "Big data" for development.

Björkegren, D., Blumenstock, J., Folajimi-Senjobi, O., Mauro, J., Nair, S.R., 2022. Instant Loans Can Lift Subjective Well-Being: A Randomized Evaluation of Digital Credit in Nigeria (No. arXiv:2202.13540). arXiv.

Björkegren, D., Blumenstock, J.E., Knight, S., 2022. (Machine) Learning What Policies Value (No. arXiv:2206.00727). arXiv.
Björkegren, D., Blumenstock, J.E., Knight, S., 2020. Manipulation-Proof Machine Learning. arXiv:2004.03865 [cs, econ].
Björkegren, D., Farronato, C., 2021. To Regulate Network-Based Platforms, Look at Their Data. Harvard Business Review. Blog post.
Björkegren, D., Grissen, D., 2019. Behavior Revealed in Mobile Phone Usage Predicts Credit Repayment. World Bank Economic Review.
Björkegren, D., Grissen, D., 2018. The Potential of Digital Credit to Bank the Poor. American Economic Association Papers and Proceedings.
Björkegren, D., Karaca, B.C., 2022. Network adoption subsidies: A digital evaluation of a rural mobile phone program in Rwanda. Journal of Development Economics 154, 102762.
Mauro, J., Björkegren, D., Blumenstock, J.. Nonparametric Causal Estimators for Multivariate Missing Data: An Application to Estimate Treatment Effects from Digital Trace Data. Working paper.
Milusheva, S., Bjorkegren, D., Viotti, L., 2021. Assessing Bias in Smartphone Mobility Estimates in Low Income Countries, in: ACM SIGCAS Conference on Computing and Sustainable Societies, COMPASS ’21. Association for Computing Machinery, New York, NY, USA, pp. 364–378.
Rolf, E., Simchowitz, M., Dean, S., Liu, L.T., Björkegren, D., Hardt, M., Blumenstock, J., 2020. Balancing Competing Objectives with Noisy Data: Score-Based Classifiers for Welfare-Aware Machine Learning. Presented at the International Conference on Machine Learning (ICML).

Other references

Athey, S., 2018. The Impact of Machine Learning on Economics. The Economics of Artificial Intelligence: An Agenda 507–547.
Bharadwaj, P., Jack, W., Suri, T., 2019. Fintech and Household Resilience to Shocks: Evidence from Digital Loans in Kenya (Working Paper No. 25604). National Bureau of Economic Research. https://doi.org/10.3386/w25604
Brailovskaya, V., Dupas, P., Robinson, J., Robinson, J., 2021. Digital Credit: Filling a hole, or digging a hole? Evidence from Malawi (Working Paper).
Bruckner, M., Scheffer, T., 2011. Stackelberg Games for Adversarial Prediction Problems, in: Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD ’11. ACM, New York, NY, USA, pp. 547–555. https://doi.org/10.1145/2020408.2020495
Donovan, K.P., Park, E., 2019. Perpetual Debt in the Silicon Savannah [WWW Document]. Boston Review. URL https://bostonreview.net/class-inequality-global-justice/kevin-p-donovan-emma-park-perpetual-debt-silicon-savannah (accessed 8.4.21).
GSMA, 2008. Taxation and the Growth of Mobile Services in Sub-Saharan Africa.
Hardt, M., Megiddo, N., Papadimitriou, C., Wootters, M., 2016. Strategic Classification, in: Proceedings of the 2016 ACM Conference on Innovations in Theoretical Computer Science, ITCS ’16. ACM, New York, NY, USA, pp. 111–122. https://doi.org/10.1145/2840728.2840730
Lucas, R.E., 1976. Econometric policy evaluation: A critique. Carnegie-Rochester Conference Series on Public Policy 1, 19–46. https://doi.org/10.1016/S0167-2231(76)80003-6
Zuboff, S., 2019. The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power, 1st edition. ed. PublicAffairs.