Research Statement

I am an empirical researcher who uses methods from statistics, machine learning, deep learning and applied microeconomics, applying both causal and predictive inference. **I focus on two broad areas of research (1) using unstructured data and predictive modeling to better assess risk in conflict prone and poor areas, and (2) understanding the economic challenges of platforms and the managerial challenges of artificial intelligence.** In the past I have focused on a mix of business and economics journals, although going forward I plan to target exclusively information systems and business journals.

I. Using Unstructured Data and Predictive Modeling to Better Assess Risk in Conflict Prone and Poor Areas

Assessing risk in developing countries is a severe impediment for firms and organizations, and in particular for those operating in the developing world. Traditional methods of measuring risk and vulnerability – using on the ground surveys – is often too expensive or of insufficient temporal resolution for making time-critical decisions. Machine learning and artificial intelligence offer unique possibilities for creating high-frequency (in some cases real-time) predictive measures of risk and vulnerability.

For much of my research in this area I use artificial intelligence (AI) computer vision models, such as Convolutional Neural Networks, applied to satellite imagery for learning economically relevant information of an area. For example, in Engstrom, Hersh, and Newhouse (2021) we apply computer vision algorithms to high-resolution satellite imagery of Sri Lanka to develop estimates of the poverty rate for local areas. We have since expanded this work to larger countries such as Mexico (Babenko, Hersh, Newhouse, Ramakrishnan, and Swartz, 2017) in a paper that was accepted at a workshop at Neural Information Processing Systems (NIPS/NeurIPS), a

leading computer science conference. In successive work, I have also explored the consequences of using closed versus openly available data for calculating official statistics. In a paper published in the journal *Information Technology for Development* (Hersh, Engstrom and Mann, 2020), we use openly available satellite images from the European Space Agency to build a poverty map in Belize and argue that data pipelines should use openly available imagery if predictive accuracy is no worse than with closed systems.

Data scarcity also affects the understanding of on the ground violence in conflict areas. In work that was published in the prestigious *Proceedings of the National Academy of Sciences* (Mueller et al., 2021), we develop a computer vision algorithm using deep learning methods to identify buildings that were destroyed from bombing attacks in Syria during their civil war. We make several methodological advances that are broadly applicable across spatio-temporal problems in computer vision networks. Data scarcity also affects the provision of credit in the consumer retail market. Partnering with the Inter-American Development Bank, I have explored how machine learning can advance the targeting of loan products for impoverished countries in Latin America (Hersh et al., 2021).

II. Understanding the Economic Challenges of Platforms and the Managerial Challenges of Artificial Intelligence

Both platforms and artificial intelligence are transforming how businesses should be managed. **In one paper (Benzel, Hersh, and Van Alstyne, 2021), currently under reject and resubmit at Management Science, we examine how application programming interfaces (APIs) influence firm performance (revenue, costs, R&D expenditures, profit) as well as the risk of cyberattacks**. APIs are the technological standard by which firms can create a structured, automatic way to transfer information within and between firms. However, the benefits and potential downsides of using this technology are not fully quantified from a managerial perspective. Using a unique

proprietary firm-level panel dataset that includes detailed information on a firm's API flow by category and type, we find that API adoption has large and positive effects for firm profit. However, we find significant increases in the probability of cyberattacks and data breach events following API adoption. This suggest that CIOs face an important tradeoff: implementing APIs decreases the costs of information transmission across a firm and can rapidly increase productivity, but also increase the risk that malicious individuals access this information.

I have recently become interested in what factors affect how managers use AI predictions, and in particular the role that explainable AI plays. In recent work **(Carter and Hersh, 2021), we implemented a field experiment at a major bank to test how managers use AI predictions and how explainable AI may affect their trust of these predictions.** We use an AI model to build predictions informing workers whether a loan they manage might encounter a delay. We embed these predictions into a dashboard and survey 690 managers to test which individual, team and model characteristics affect their trust of the AI predictions. We randomly assign half of the managers to receive an "explainable AI" treatment that include additional dashboards showing how the model made its prediction. We find mixed evidence that explainable AI affects trust overall, and further find that senior managers and self-reported machine learning novices are least likely to trust AI. However, these AI reluctant groups are most likely to respond positively to explainable AI, increasing their trust of AI by 5-7 times if they receive an explainable AI module. We are currently drafting the paper in preparation for submission to Management Science.

Another strain of research examines the effectiveness and efficacy of online media anti-piracy policies. In one paper (Danaher, Hersh, Smith, and Telang, 2020), published in *MIS Quarterly*, we analyze how supply-side piracy interventions may affect legal and illegal channels of consumer online media streaming. We analyze three court-ordered website blocking events of increased effectiveness, affecting consumers in the UK. We first study Internet Service Providers' blocking of 53 major piracy sites in 2014 and we then study two smaller waves of blocking – the blocking of 19 piracy sites in 2013 and the blocking of The Pirate Bay in 2012. We show that blocking 53 sites in 2014 caused

treated users to decrease piracy and to increase their usage of legal subscription sites by 7-12%. In another paper with Brett Danaher (Danaher and Hersh, 2020), currently under review, we revisits how policy-makers should think about anti-piracy enforcement. Using panel data on streaming video users before and after a major link piracy site was shutdown, we examine the types of users who continue to pirate even after a large enforcement crackdown, and those who move to more legitimate streaming services. We find that while total piracy goes down, demographics do not play a large role in who continues to pirate. Income, however, is highly predictive of legal consumption use, suggesting that pricing behavior of alternative streaming services matters more than criminal intent.

In one paper (Lang, Lang and Hersh, 2020) we examine the relationship between the growth of smartphones and traffic accidents in California between 2001 and 2011. We link cellular coverage along a highway in 2016 to the location of antenna towers and then apply machine learning techniques to predict coverage between 2001 and 2011. We find that car accident rates increase when 3G cell phone coverage is introduced, even when controlling for traffic volume. This work has a revise and resubmit at *Journal of Economic Behavior and Organization*.

References

Babenko B., Hersh J., Newhouse D., Ramakrishnan A., and Swartz T.. Poverty Mapping Using Convolutional Neural Networks Trained on High and Medium Resolution Satellite Images, With an Application in Mexico, Neural Information Processing Systems (NIPS 2017)

Benzell, S., Hersh J. Lagarda Cuervas, G., Van Alstyne M.V. (2021). How APIs Create Growth by Inverting the Firm. *1st round revisions submitted,* **Management Science**

Danaher B., and Hersh J. (2020). Whom Should Antipiarcy Enforcement Actions Target? Evidence from the MegafilmesHD Shutdown in Brazil, *Under review*

Danaher B, Hersh J, Smith M, and Telang R, The Effect of Piracy Website Blocking on Consumer Behavior, (2020), **MIS Quarterly**

Engstrom R., Hersh J., Newhouse D., (2021). Poverty from Space: Using High Resolution Satellite Imagery for Estimating Economic Well-being, **The World Bank Economic Review***; lhab015,*

Hersh J., Lang B., and Lang M. (2020). Digitally Distracted at the Wheel: Myopic Smartphone Use and Car Accidents, *Revisions requested,* **Journal of Economic Behavior and Organization**

Hersh J., Engstrom R., Mann M. Open Data for Algorithms: Mapping Poverty in Belize Using Openly Available Satellite Features and Machine Learning. **Information Technology for Development**, 1-30.

Hersh J. Carter S., (2021). Will AI Accelerate the Intra-Firm Digital Divide? Evidence from a Field Experiment on How Managers use Explainable AI

Jonathan Hersh, Lucia Martin Rivero, & Janelle Leslie. Where are the unbanked in Belize? Using Machine Learning small area estimation to improve Financial Inclusion geographic targeting. Inter-American Development Bank Monograph, May 2021

Mueller M., Groger A., Hersh J., Matranga A., Serrat J. Analyzing Conflict From Space: Identification of Physical Destruction During the Syrian Civil War with Artificial Intelligence . **Proceedings of the National Academy of Sciences.** June 8, 2021 118 (23) e2025400118;