Research Statement

Jacopo Perego
Columbia Business School
Associate Professor of Economics
email: jacopo.perego@columbia.edu
web: www.jperego.com

September 2022

In my research, I use both theoretical and experimental methods to investigate how information affects our societies. From the products we buy to the political candidates we support, I study how information became a key productive input for many firms in the digital economy. My interest in this topic is motivated by the recent development of information technologies, which fueled trillion-dollar industries and created many new challenges for regulators and policymakers.

Broadly speaking, my research contributes to the field of information economics. More specifically, it focuses on two main themes:

A. Information Design *My work on this: [1], [2], [4], [7]*

I study how to design information in order to influence the behavior of strategic agents—for example, via rating systems and product recommendations. Using both theory and experiments, I study how the different information environments in which agents interact can ultimately affect their behavior. To goal of this research is both methodological—as it advances our ability to analyze and solve these design problems—as well as practical—as it can provide guidance on what features these systems should have in practice.

B. Information Markets *My work on this: [3], [5], [6], [8], [9]*

I study markets in which information is traded, either as a productive input (e.g., the market for personal data) or as a final good (e.g., the news media industry). My research identifies externalities that riddle these markets. Its goal is to advance our understanding of these prominent markets and, ultimately, inform regulators on how to minimize their inefficiencies.

The rest of this document briefly describes my work and clarifies how my papers are connected together.

A. The Design of Information

The first set of papers contribute to the fast-growing literature on information design. This is a paradigm that studies how to optimally persuade interacting agents exclusively via the provision of information. Information design complements the more classical approach of *mechanism* design, in which the allocation of information is exogenous but the designer can affect behavior by changing agents' incentives (e.g., with transfers).

Mathevet, Perego, and Taneva (2020, "On Information Design in Games," Journal of Political Economy) contributed to the foundation of this new paradigm. It is one of the early papers in this area that studied the problem of how to optimally influence a group of agents who interact strategically (i.e., in a "game"). Most of the previous literature focused on how to persuade a single agent who acts in isolation. A theory of information design in games is valuable for two reasons. The first one is practical. Many important applications feature multiple interacting agents, rather than a single one. For example, an online marketplace may provide information to competing sellers; a startup may provide information to a group of early investors, etc. The second reason is conceptual. When multiple agents interact, the provision of public information may no longer be optimal. In these cases, the designer may need to provide information privately, which requires harnessing complex hierarchies of higher-order beliefs. Our paper tackles this problem head-on. We provide tools that help researchers better understand and solve these kinds of information design problems. We show that any problem can be solved with a two-step approach: the first step optimizes among "simple" distributions only, and the second step optimally combines these distributions subject to Bayesian plausibility. The latter step generalizes the concavification approach of Kamenica and Gentzkow (2011). Incidentally, the tools provided by this paper allow us to analyze information design problems with a variety of solution concepts and equilibrium selection rules.

The information-design paradigm described above rests on a number of important assumptions. For example, it assumes the designer can choose any information structure, without constraints. However, when using information to influence agents' behavior, one kind of constraint seems especially important: Agents could talk with each other. Galperti and Perego (2020, "Information Systems," Working Paper) is motivated by this simple observation. We relax one of the main tenets of the theory of information design by assuming that agents share some of the information they receive from the designer. We do so by providing a new revelation-principle result that allows us to analyze the properties of this class of constrained information design problems. Our leading application is in organizational economics. Consider a CEO who can either organize regular meetings between two divisions of a firm (e.g., R&D and marketing) or institute a firewall that prevents information to leak across these two divisions. The question of how to optimally harness the information flows inside an organization is a classic one in the organization-economics literature. Most papers in this literature assume a given information structure. In practice, decisions about the organizational structure are made without knowing what information these divisions will eventually obtain. Our methods allow to revisit this question a provide a robust answer that makes minimal assumptions on the initial allocation of information.

The information-design paradigm rests on another key assumption, namely that senders can fully commit to an information structure. The importance of this assumption is at the center of

Frechette, Lizzeri, and Perego (Forthcoming, "Rules and Commitment in Communication: An Experimental Analysis," Econometrica), where we study information provision under *partial* commitment. Our framework nests together models of cheap talk, information disclosure, and Bayesian persuasion. This framework predicts that when information is unverifiable, a sender with more commitment power transmits more information to the receiver. In contrast, when information is verifiable, a sender with more commitment power transmits less information to the receiver. We leverage these contrasting forces to design an experiment that allows us to comprehensively test how senders use commitment in communication and how receivers react to it. Our main results find strong treatment effects in line with the main qualitative predictions of the theory. We also uncover important quantitative deviations from the theory. Specifically, we find that a form of commitment "blindness" leads some senders to persistently over-communicate when information is verifiable and under-communicate when it is not. This generates an unpredicted gap in information transmission across the two rules, suggesting a novel role for verifiable information in practice.

In Farina, Frechette, Lizzeri, and Perego (2022, "Verifiability in Communication: An Experimental Analysis," Work in Progress), we drop the commitment assumption altogether and study a very classic question: How does verifiability affect the communication between a sender and a receiver? This question has been studied extensively, both theoretically and experimentally. In theory, we know that, under some conditions, verifiability leads to full disclosure. In the data, disclosure is often incomplete. It is well-known that, in practice, people do not always play Nash equilibria (or suitable refinements). Is this failure of the unraveling principle so severe to invalidate our theories of communication? Our experiment is designed to answer this question. We develop a theory of *partial* verifiability that allows us to introduce variation in the degree to which information is verifiabile. This leads to a rich set of comparative statics, which we bring to the lab as a way to test key qualitative prediction of the theory. This project is currently in the data-collection phase.

B. Markets for Information

The second set of papers is more applied and studies markets in which the principal good that is traded is information, either as a productive input (e.g., the market for personal data) or as a final good (e.g., the news media industry).

Perego and Yuksel (2022, "Media Competition and Social Disagreement," Econometrica) analyzed the market for political information—the news-media industry. We study firms that compete to sell information to voters. Our main result highlights a novel equilibrium channel through which a more competitive market—specifically, a market where more firms compete—decreases the efficiency of electoral outcomes. Competition leads firms to specialize their informational products. Crucially, they do so by providing relatively less information on issues

that are of common interest and relatively more information on issues on which agents' preferences are heterogeneous. This enables agents to acquire information about different aspects of the policy, specifically, those that are particularly important to them. This leads to an increase in social disagreement, which has negative welfare implications. We establish that, in large enough societies, competition makes every agent worse off by decreasing the utility that she derives from the electoral outcome. Furthermore, we show that this decline cannot be compensated by the decrease in prices resulting from increased competition.

This paper contributes to a large industrial-organization literature on spatial competition by innovating in three ways. First, we explicitly model equilibrium interactions between vertical and horizontal competition. Thanks to this, we can show that competition leads firms to disinvest from vertical features—which are beneficial to all consumers—in favor of horizontal features—which are beneficial only to a niche segment. Second, we study the consequences of specialization in a context in which private consumption generates social externalities (in our case through voting). This feature is common to many markets, well beyond our political-economy application. Third, our firms sell information and, to account for this, we build Bayesian foundations into a spatial competition model. For these reasons, we hope that this model can become a workhorse to study questions at the intersection of political economy and industrial organization.

In a series of new papers, my coauthors and I study the data-brokerage industry. This research agenda was awarded a multi-year research grant from the National Science Foundation (SES-2149315 and SES-2149289).

The first of these papers is Galperti, Levkun, and Perego (2022, "The Value of Data Records"), a working paper that is currently *R&R* at the Review of Economics Studies. It is generally understood that personal data contributes to fueling multi-billion-dollar industries. But how much of this total value is created by the data of each single individual? This basic question is at the core of some of the recent debates regarding the future of data markets: how to design them to compensate individuals for their data, how to conduct demand analysis for data brokers, and to what extent data is a source of market power. Yet, the value of personal data is not well understood and can be difficult to assess, as illustrated by our opening example in the Introduction. Our paper proposes a unifying and inherently classical approach, whose roots are in information design. Our analysis uncovers a novel externality between data records. This arises not from exogenous correlation between records, but endogenously from how the intermediary uses records to withhold information from the agents. As such, this externality is a hallmark of intermediation problems due to their inherent conflicts of interest between the various parties involved. Our analysis has practical implications. First, we show that our values constitute a useful benchmark for compensating individuals for their data and discuss its desirable features. Second, we draw an analogy between how an intermediary values data records and

a consumer from standard consumer theory. This opens the door to applying well-known analytical tools for studying the demand side of data markets and optimal strategies to invest in data acquisition. Third, we study the intermediary's WTP for better records and analyze both the extensive and intensive margin of refining a database.

The second paper in this agenda is Galperti and Perego (2022, "The competitive price of data records"), which we are preparing for the American Economic Association, *P&P*. In this paper, we make precise the sense in which the values characterized by Galperti, Levkun, and Perego (2022) constitute a useful benchmark. Using standard linear-programming arguments, we find that these values corresponds to the equilibrium *prices* in a competitive market where a large number of data intermediaries compete to acquire buyers' data records and use them in a revenue-maximizing way.

The third and last paper in this agenda on data markets is Galperti and Perego (2022, "Privacy and the Value of Data"). In this work, we build on our framework and ask how privacy regulation affects what data are collected, how they are used, and their economic value? Privacy regulation constrains the data intermediary problem by granting individuals more control over how their personal data are used. We can embed several of these constraints in the mathematical formulation of our framework—for instance, whether a platform can refuse to serve a customer who does not share her data or is required to provide her with at least a basic service. We find that, in addition to the direct payoff it generates for the intermediary, the value of an individual's data record may reflect externalities that arise because its use helps convincing other individuals to share their data. One insights that emerges from our preliminary analysis is that privacy regulations may not only shift wealth from data-users to data-sources (i.e., from platforms to consumers), but also change the values of data records themselves, thus redistributing wealth across consumers.

Finally, Perego and Yuksel (2017, "Searching for Information and the Diffusion of Knowledge") study an informal market for information that is akin to a social media. The paper proposes a dynamic learning model in which an agent can either produce information for himself (work) or learn from others by observing their past behavior. This creates a fundamental trade-off between the production of new information and the aggregation of information created by others. Both activities are socially valuable. Agents are heterogeneous in two dimensions. First, more connected agents can observe past behavior more accurately and exert a stronger influence on the behavior that is observed by future generations. Second, players have idiosyncratic biases. We allow for uncertainty about the average bias of the society. We show that the structure of social connections interacts with this type of uncertainty affecting equilibrium information production and its social diffusion in ways that are complex and subtle. We identify two main channels through which players generate externalities on each other. The first, *the learning effect*, captures the intuitive idea that more connected players, who have ac-

cess to more information, take more informative actions and hence increase the attractiveness of learning from others for the rest of the society. The second, *the amplification effect*, refers to an equilibrium mechanism by which uncertainty about others' preferences reverberates and amplifies throughout the society as more people choose to learn from others. We show that increasing the connectivity of the society, or generally shifting influence towards more connected types, can lead to a strict decrease in the quality of social information. Moreover, we find that the socially optimal allocation into learning activities can differ dramatically from the equilibrium one: under certain conditions, the planner would reverse the equilibrium allocation, forcing highly connected players to work, and moderately connected ones to learn from others.

References

[1] Farina, A., G. Frechette, A. Lizzeri, and J. Perego (2022): "Verifiability in Communication: An Experimental Analysis," *Work in Progress*.

[2] Frechette, G., A. Lizzeri, and J. Perego (Forthcoming): "Rules and Commitment in Communication: An Experimental Analysis," ***Econometrica***.

[3] Galperti, S., A. Levkun, and J. Perego (2022): "The Value of Data Records," *Working Paper*, currently R&R at ***Review of Economics Studies***.

[4] Galperti, S. and J. Perego (2020): "Information Systems," *Working Paper*.

[5] ——— (2022): "The Competitive Price of Data Records," *Work in Progress*, manuscript in preparation for ***American Economic Association, P&P***.

[6] ——— (2022): "The Value of Data and Privacy," *Work in Progress*.

[7] Mathevet, L., J. Perego, and I. Taneva (2020): "On Information Design in Games," ***Journal of Political Economy***, Vol 128-4.

[8] Perego, J. and S. Yuksel (2017): "Searching for Information and the Diffusion of Knowledge," *Working Paper*.

[9] ——— (2022): "Media Competition and Social Disagreement," ***Econometrica***, Vol 90-1.