RESEARCH STATEMENT

VASILIS SYRGKANIS

My research lies at the intersection of theoretical computer science, machine learning and economics. It aims at understanding the *foundations* of large scale computer systems where *strategic* users interact. I strive for data-driven methodologies that utilize the *large availability of data* in these systems, methodologies which I have applied to real world data-sets from online auction marketplaces. My work combines and develops tools and techniques in a variety of areas such as algorithms, game theory, machine learning, mechanism design, econometrics, data science, and computational complexity.

The internet has provided an unprecedented platform for an exponentially growing digital economy. Designing these modern marketplaces requires the development of theoretical and empirical tools at the intersection of theoretical computer science and economics. It is not sufficient to develop efficient algorithms that given an input produce quickly some desirable output. The input is owned by different selfish entities, who will try to manipulate it so that the algorithmic output changes to their own advantage. Thus we need a theory that addresses the *equilibrium* properties of *algorithms* when the input is strategically chosen. On the other hand, when designing economic mechanisms and marketplaces we need to address computational concerns, both from the perspective of the market designer and from the perspective of the market users. The latter desideratum takes two incarnations: (i). when addressing the *equilibrium* properties of markets we need *efficiently computable* notions of equilibrium, i.e., equilibria that users can arrive at by invoking computationally efficient simple algorithms; (ii). the algorithmic part of a mechanism that given some input, produces an allocation of market resources (goods on eBay, advertisement slots, computational resources in the cloud) has to have a very *fast* and ideally *distributed* implementation.

My *theoretical* work addresses exactly these considerations. In a series of papers [3, 5, 6, 9, 11, 13, 14, 21, 24, 25], we have provided a unifying theory for the equilibrium analysis of the quality of outcomes in strategic environments, in a manner that is robust to behavioral and information assumptions from the perspective of the participants (see Section 1). This line of work culminated in a co-authored survey [20], invited to the Journal of Artificial Intelligence Research. My research offers a new analysis tool to the economics literature, through the lens of *approximation*. Prior to our work the equilibrium analysis of auctions required many restrictions: one item being sold, or complete information or strong symmetry assumptions on the participants or the goods. Extending beyond these settings was hindered by the inability to fully characterize equilibrium outcomes. Our approach bypasses this hindrance, by directly providing properties of any equilibrium and even non-equilibrium learning concepts, without the need of an analytic characterization.

My *empirical and data-driven* work applies approaches from the algorithmic and machine learning world to address econometric problems when analyzing data-sets from strategic environments [17, 10] (see Section 2), and learning theory problems in the presence of incentives and competition [2, 15, 22, 16] (see Section 3). One key question that I have addressed in a series of papers [17, 10] is how to infer unobserved parameters from observed strategic interactions, under the assumption that players use learning algorithms rather than that they have arrived at the classic economic equilibrium. Empirical evidence from sponsored search auction markets suggests that players are not playing according to a static equilibrium. We posit that the observed behavior is the outcome of some form of online learning and based on this assumption propose an econometric method for inferring the private parameters of the participants or for inferring directly the unobserved quality of the observed allocation of resources. I have applied these methods to data-sets from Microsoft's sponsored search auction marketplace. This work [17] received the Best Paper Award at the ACM conference on economics and computation. In another work [22], which received a Best Paper Award at the NIPS conference, we examine the convergence properties of learning algorithms when applied to game theoretic environments.

Microsoft Research, NYC, October 2015

1. Efficiency in complex markets

The strategic decisions that players face in digital marketplaces are rarely simple. Typically players participate in multiple substitute markets at the same time, they rarely have complete knowledge of the game that they are playing, but rather only know what options are available to them (e.g. submit a bid per click in a sponsored search auction setting), while the environment and competition is dynamically changing in an unpredictable manner. In many of these settings it is not even in the hand of the designer to turn the game into a simpler one where all that the market participant needs to do is report his needs truthfully and let the market mechanism handle the optimization for him. Such *truthful* designs typically assume that all resources are handled by the same market maker, that they are all available at the same time and finally, they tend to lead to unintuitive or impractical rules, which the participants rarely trust.

The latter realization renders necessary the *equilibrium analysis* of simple and non-truthful market mechanisms. Classical game theory assumes that in such a strategic interaction the players will pick strategies that constitute a mutual best-response, i.e., a Nash equilibrium. Given the computational complexity of computing such a set of mutual best-responses and the information required from the players, it is hard to envision that participants will reach such a notion of equilibrium. Moreover, given that in most settings arising in electronic markets, players repeatedly participate in the same game and the cost of experimentation is rather small, it is more reasonable to believe that they will use some form of simple adaptive learning algorithm to identify how to play optimally. One major line of my work [11, 24, 21, 25, 9, 5, 13, 3, 6, 14] analyzes the efficiency of strategic outcomes, commonly referred to as the *price of anarchy*, in such complex markets. Below I describe some of the key points of this line of my work.

Efficiency under learning behavior and incomplete information. In [21, 25, 9, 14] we identify a sufficient property for worst-case efficiency guarantees at equilibrium in games and mechanisms. The property, dubbed *smoothness* of the game or mechanism, was defined in the context of complete information games by Roughgarden [19] and was adapted to mechanisms in my co-authored work [25]. It is a requirement for the game to admit good strategies for each player through which they can produce a good fraction of their contribution to the optimal welfare by unilaterally deviating, without incurring a high cost relative to the revenue of the current solution. We have shown [25] that many simple mechanisms and games satisfy this property, in a diverse set of strategic environments such as bandwidth allocation and combinatorial auctions. Most importantly, these efficiency guarantees are very robust as they hold even if:

i. players' parameters, such as valuations, are drawn privately and stochastically [21];
ii. the game is not at equilibrium and players use adaptive learning strategies that satisfy the no-regret assumption, widely studied in online learning theory, [9];
iii. the population playing the game is constantly and almost adversarially changing with high frequency [14].

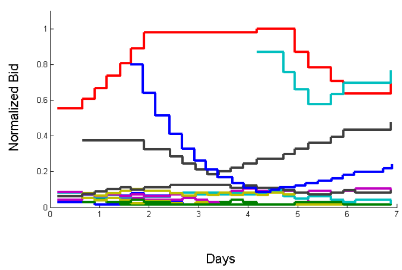
Interestingly, the last result [14], provides a novel connection between price of anarchy analysis and *differential privacy* [4].

Further topics: composability, algorithmic characterizations, large markets. This line of work uncovers several other important properties of smooth mechanisms. In [25], we initiate the study of local-to-global guarantees in mechanism design and show that smoothness is a composable property: smoothness locally at each mechanism implies global market efficiency, even if players participate in multiple mechanisms simultaneously and under some complement-free assumption on the resources sold by different mechanisms. In [13], we address algorithmic characterizations of smooth mechanisms and show that a sufficient condition is that the algorithm that is used for the allocation of resources can be viewed as running a greedy algorithm in a simple feasible allocation space (a matroid). In [6] we identify sufficient conditions for a smooth mechanism or game to become more efficient as the market grows large and each individual player has diminishing impact on the outcome.

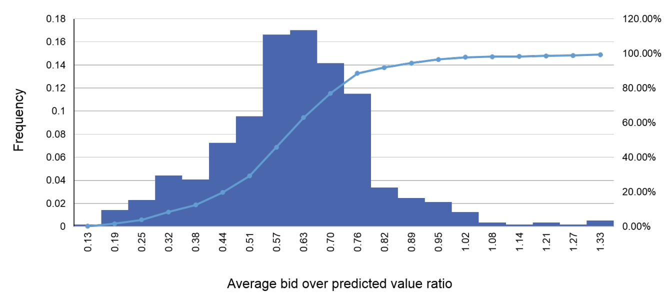
2. Learning and data science in strategic environments

Another main goal of my research is to utilize the large availability of data in modern computer systems and markets so as to understand and improve their performance. While classical statistical learning theory copes mostly with settings where the input data are noisy, in environments where strategic interactions are at play, the data are not only noisy, but are also strategically chosen by selfish entities.

For instance, in online auction markets, though we observe the actions of participants, such as the bids of an advertiser in a sequence of auctions, we never observe his true value for appearing in an advertisement slot, or his value for getting a click or in general the objective that he is trying to optimize. A question of major importance is how to estimate these parameters from the observed behavior. This question is at the center of econometric theory. However, most of econometrics assumes that the system has reached an equilibrium. Stable equilibrium behavior is refuted from data-sets. As we depict in Figure 1a, a single advertiser might change his bids many times during the period of a week.



(a) Bidding data from `BingAds` over a period of a week.



(b) Valuation inference analysis applied to advertiser in Microsoft's sponsored search auction system.

Econometrics for learning agents. How does the theory change if instead we observe the market in a transient, non-equilibrium state? In [17], we propose an econometric method for inferring player private valuations in an auction setting assuming that the observed behavior is the outcome of a learning algorithm that satisfies the *no-regret assumption*, i.e. the average utility of the algorithm is at least as good as the best fixed strategy in hindsight, less an error term that vanishes to zero over time. The no-regret learning assumption is a good fit for online auction environments for several reasons: (i). many simple, multi-purpose learning algorithms satisfy the assumption, such as *multiplicative weight updates* [12, 7] or *regret matching* [8]; (ii). these algorithms only require black-box utility feedback from the action that the player used and do not require any other knowledge of how the auction works or what the competition is, iii) the benchmark of beating a fixed strategy over time seems a minimal requirement for any reasonable learning algorithm.

Apart from capturing learning behavior our method is also very efficient from a computational and sample-complexity viewpoint and easily applicable to data-sets from auction markets. We have applied our method to data-sets from Microsoft's sponsored search auction system leading to inference of how much players shade their valuation in the auction. Figure 1b depicts an example of such a result when applying our method to the listings of an advertiser on `BingAds`. We find that typically advertisers bid, on average, approximately 60% of their value.

Welfare guarantees from data. In [10], we address the task of inferring welfare guarantees from data, rather than exact player valuations. We show that, as an easier task, it admits a simpler econometric method. This work bridges the gap between worst-case welfare guarantees produced by my theoretical research and average-case instances that arise in practice.

3. Machine learning theory in strategic environments

Machine learning algorithms are currently deployed in a variety of strategic environments. For instance, algorithms used in bid optimization tools for online ad auctions are competing against other instances of similar learning algorithms. In other settings, such as recommendation systems, the feedback of a learning algorithm comes from actions taken by strategic individuals trying to maximize their own utility. When analyzing properties of these algorithms, we should be thinking about competition and incentives. The results described Section 1 can be viewed as one instance of analyzing the efficiency properties of outcomes of machine learning algorithms when competing with each other. My research addresses several topics other than economic efficiency.

Fast convergence of learning algorithms in games. Learning algorithms are typically analyzed in worst case, adversarial environments. In [22] we analyze properties of online learning algorithms when deployed in a game theoretic setting, alongside other learning algorithms. We show that if these algorithms are relatively stable in their predictions between time-steps and have some form of recency bias, then their learning rates are much faster than if they were deployed in completely adversarial environments. An example is the large class of *follow-the-regularized-leader* rules modified so as to put more weight on recent observations. Our work extends previous work of [1, 18], which analyzed only two-player zero-sum games.

Computational complexity of online learning. Out-of-the-box no-regret learning algorithms converge fast and in polynomial time when the number of actions available to each player is explicitly given or is polynomial in the description of the game. In some auction settings, the strategy space is exponential. In [2] we consider the computational complexity of no-regret algorithms in such environments and show that there cannot exist any efficient no-regret algorithm (subject to standard computational assumptions). We propose alternative notions of adaptive game-playing that admit efficient implementations in such settings, founded on the notion of a Walrasian market equilibrium. In another direction [3], we design simple combinatorial auctions where the number of actions available to each player is small enough that existing online learning algorithms are efficient.

Incentive compatible learning. In [15, 16] we design incentive-compatible learning algorithms in exploration-vs-exploitation learning settings, typically referred to as *bandit learning*, with asymptotically optimal learning rates. Such settings arise in crowdsourcing environments where a recommendation system is trying to collect information from its users about the available options, i.e. *explore*, so as to make more informed decisions for future users, i.e. *exploit*. The learning algorithm used in the information collection stage, must be designed appropriately so that users would want to follow the recommendation, which is formalized as an *incentive* constraint.

4. Research agenda

My goal is to combine theoretical work at the intersection of theoretical computer science and economics, together with empirical analysis of application domains of interest, such as electronic marketplaces. Advancing the state of the art in our theoretical understanding of problems involving computation, incentives and limited information, is quintessential to addressing the problems that arise in modern large scale computer systems and modern large scale markets. On the other hand, we need to take advantage of the plethora of data-sets available to inform our theoretical models and make our theories more practically relevant and with higher immediate impact.

At the risk of being too concrete, let me first give an example related to my ongoing work, portraying this interplay between theory and data. Most auction models in sponsored search auctions assume that the clickability of an ad is a product of its quality factor and a factor related to the position it acquired. This is hardly the case in practice. However, it is not fruitful to simply assume an ad-hoc more complex model and then do theoretical analysis. We can use data-sets from sponsored search markets to understand what is the right theoretical model to adopt. In an ongoing work I am empirically testing such models in Microsoft's sponsored search auction system, with the goal that the empirical analysis will also lead to subsequent novel theoretical research in auction design.

Theoretical research. Several theoretical directions in the Econ-CS field remain largely open. Even though we have a very good understanding of efficiency of complex markets in terms of their welfare, there are several other important *quality metrics* which have not yielded to our current methodologies, such as fairness of the allocation or revenue of the market maker. Designing marketplaces for objectives other than welfare and particularly when participants will not play at equilibrium, but rather use learning algorithms, is largely open.

My research will also continue to explore questions related to the learning and dynamic optimization problems that both the market makers and the participants are facing. How does one use data to learn optimal prices, or optimal auction schemes, or in general optimize the design of her marketplace? How does a player facing a set of decisions in a market behave optimally or learn how to behave optimally over time? There is a host of dynamic optimization problems in this area that do not yet have efficient algorithmic solutions, especially when taking into account competition and incentives.

Another direction I plan to pursue stems from the realization that information is one of the most valuable commodities in modern markets. Several data-markets exist for the selling of information and several marketplaces, such as ride-sharing or job market platforms, are designed to incent users via information revelation schemes. Viewing information both as a commodity and as a design tool opens a lot of theoretical questions at the intersection of algorithms and mechanism design. One example of my work on the topic [23] addresses the role of information in auctions, motivated by settings in targeted advertising.

Empirical and data-driven research. My work on econometrics for learning agents opens a very promising research area at the intersection of online learning, machine learning and econometrics. I plan to pursue further such connections, both from the perspective of modeling players as machine learning agents, as well as using machine learning techniques for inferring complex econometric models. On a concrete level, one of my main goals is to understand how players behave in complex markets. What is the most predictive model of learning behavior? Analyzing such predictive models of behavior is quintessential to addressing market optimization questions; we need a way of predicting the reaction of strategic agents to changes in the market design. In the absence of an equilibrium assumption, we need alternative models of dynamic player behavior. I expect that my strong ties with Microsoft Research will allow me to continue testing econometric theories on real-world data sets.

Application domains. There are several new contexts in which we can apply both our theoretical and empirical methodologies. In recent years we have seen the rise of platform economies (e.g., ad auctions, sharing economies, job market services, crowdsourcing services, cloud computing services, internet-of-things), where the system merely offers a platform for strategic agents to interact. The two-sided nature of these economies opens up new theoretical and empirical questions that have not been well-addressed. Another important application domain is that of energy markets. The smart-grid infrastructure allows for energy to be dynamically procured via complex auction schemes. Understanding the equilibrium properties of these auctions is of great importance for designing efficient energy marketplaces.

In sum, my research plan is to address fundamental theoretical questions at the intersection of economics and computer science, as well as explore the new opportunities that arise from the large availability of data in the digital economy, with the goal of having immediate impact on the design of modern marketplaces.

References

[1] Constantinos Daskalakis, Alan Deckelbaum, and Anthony Kim. Near-optimal no-regret algorithms for zero-sum games. *Games and Economic Behavior*, 92:327 – 348, 2015.

[2] Constantinos Daskalakis and **Vasilis Syrgkanis**. Learning in auctions: Regret is hard, envy is easy. *working paper*, 2016.

[3] Nikhil R. Devanur, Jamie Morgenstern, **Vasilis Syrgkanis**, and S. Matthew Weinberg. Simple auctions with simple strategies. In *Proceedings of the Sixteenth ACM Conference on Economics and Computation, EC '15, Portland, OR, USA, June 15-19, 2015*, pages 305–322, 2015.

[4] Cynthia Dwork and Aaron Roth. The algorithmic foundations of differential privacy. *Found. Trends Theor. Comput. Sci.*, 9(3-4):211–407, August 2014.

[5] Uriel Feige, Michal Feldman, Nicole Immorlica, Rani Izsak, Brendan Lucier, and **Vasilis Syrgkanis**. A unifying hierarchy of valuations with complements and substitutes. In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence, January 25-30, 2015, Austin, Texas, USA.*, pages 872–878, 2015.

[6] Michal Feldman, Nicole Immorlica, Brendan Lucier, Tim Roughgarden, and **Vasilis Syrgkanis**. The price of anarchy in large games. In *Proceedings of the Forty-Eighth Annual Symposium on the Theory of Computing*, STOC'16, 2016.

[7] Yoav Freund and Robert E Schapire. Adaptive game playing using multiplicative weights. *Games and Economic Behavior*, 29(1):79–103, 1999.

[8] Sergiu Hart and Andreu Mas-Colell. A simple adaptive procedure leading to correlated equilibrium. *Econometrica*, 68(5):1127–1150, 2000.

[9] Jason D. Hartline, **Vasilis Syrgkanis**, and Éva Tardos. No-regret learning in repeated bayesian games. In *Proceedings of the Twenty-ninth Annual Conference on Neural Information Processing Systems, NIPS'15, Montreal, Canada, Dec. 7-12, 2015*, 2015.

[10] Darrell Hoy, Denis Nekipelov, and **Vasilis Syrgkanis**. Robust data-driven efficiency guarantees in auctions. *CoRR*, abs/1505.00437, 2015.

[11] Renato Paes Leme, **Vasilis Syrgkanis**, and Éva Tardos. Sequential auctions and externalities. In *Proceedings of the Twenty-Third Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2012, Kyoto, Japan, January 17-19, 2012*, pages 869–886, 2012.

[12] Nick Littlestone and Manfred K Warmuth. The weighted majority algorithm. *Information and computation*, 108(2):212–261, 1994.

[13] Brendan Lucier and **Vasilis Syrgkanis**. Greedy algorithms make efficient mechanisms. In *Proceedings of the Sixteenth ACM Conference on Economics and Computation, EC '15, Portland, OR, USA, June 15-19, 2015*, pages 221–238, 2015.

[14] Thodoris Lykouris, **Vasilis Syrgkanis**, and Éva Tardos. Learning and efficiency in games with dynamic population. In *Proceedings of the Twenty-Seventh Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2016, Arlington, VA, USA, January 10-12, 2016*, 2016.

[15] Yishay Mansour, Aleksandrs Slivkins, and **Vasilis Syrgkanis**. Bayesian incentive-compatible bandit exploration. In *Proceedings of the Sixteenth ACM Conference on Economics and Computation, EC '15, Portland, OR, USA, June 15-19, 2015*, pages 565–582, 2015.

[16] Yishay Mansour, Aleksandrs Slivkins, **Vasilis Syrgkanis**, and Steven Wu. Bayesian incentive compatible learning in games. *working paper*, 2015.

[17] Denis Nekipelov, **Vasilis Syrgkanis**, and Éva Tardos. Econometrics for learning agents. In *Proceedings of the Sixteenth ACM Conference on Economics and Computation, EC '15, Portland, OR, USA, June 15-19, 2015* ***(Best Paper Award)***, pages 1–18, 2015.

[18] Alexander Rakhlin and Karthik Sridharan. Optimization, learning, and games with predictable sequences. In *NIPS*, pages 3066–3074, 2013.

[19] T. Roughgarden. Intrinsic robustness of the price of anarchy. In *Proceedings of the 41st annual ACM symposium on Theory of computing*, STOC '09, pages 513–522, New York, NY, USA, 2009. ACM.

[20] Tim Roughgarden, **Vasilis Syrgkanis**, and Éva Tardos. Price of anarchy in auctions. *working paper, invited to the Journal of Artificial Intelligence*, 2015.

[21] **Vasilis Syrgkanis**. Bayesian games and the smoothness framework. *CoRR*, abs/1203.5155, 2012.

[22] **Vasilis Syrgkanis**, Alekh Agarwal, Haipeng Luo, and Robert E. Schapire. Fast convergence of regularized learning in games. In *Proceedings of the Twenty-ninth Annual Conference on Neural Information Processing Systems, NIPS'15, Montreal, Canada, Dec. 7-12, 2015* ***(Best Paper Award)***, 2015.

[23] **Vasilis Syrgkanis**, David Kempe, and Éva Tardos. Information asymmetries in common-value auctions with discrete signals. In *Proceedings of the Sixteenth ACM Conference on Economics and Computation, EC '15, Portland, OR, USA, June 15-19, 2015*, page 303, 2015.

[24] **Vasilis Syrgkanis** and Éva Tardos. Bayesian sequential auctions. In *ACM Conference on Electronic Commerce, EC '12, Valencia, Spain, June 4-8, 2012*, pages 929–944, 2012.

[25] **Vasilis Syrgkanis** and Éva Tardos. Composable and efficient mechanisms. In *Symposium on Theory of Computing Conference, STOC'13, Palo Alto, CA, USA, June 1-4, 2013*, pages 211–220, 2013.