Research Statement
Matt Clancy

My graduate research has two strands both centered on the economics of innovation. The primary strand is a new model of knowledge production that combines combinatorial theories of innovation with Bayesian learning about the underlying structure of knowledge creation. The second strand examines the impact of different technological policies, such as a carbon tax and a mandate, on the kinds of innovation attempted by the private sector. Going forward, I plan to continue working in these and related areas.

The first two papers of my dissertation explore the implications of my new model of knowledge production. In my model, researchers have access to a set of primitive knowledge elements that can be combined to form ideas, where a new combination is a new idea. Underlying parameters governing the connections between elements stochastically determine whether a given combination yields a useful idea. These underlying parameters are unknown to researchers, and I represent their beliefs about these parameters with a prior distribution. As researchers attempt to combine elements and create ideas, they observe signals about these underlying parameters, which they use to improve their beliefs via Bayesian updating. While research always remains a stochastic process, researchers do learn which elements are relatively likely to work together and which are relatively unlikely to work together.

I embed this production function into a very simple model of research incentives, where a single forward looking researcher receives a reward for discovering new and useful combinations, but pays a cost to conduct research. The researcher’s problem is to determine what combinations of elements to try and combine. I use a combination of special cases amenable to analytic solutions and computer simulations (written in python) to derive characteristics of the optimal research strategy and dynamics in research productivity.

This work suggests a number of characteristics that help predict when a pair of elements is likely to be combined in research projects. For example, pairs of elements that have successfully yielded useful ideas in the past are more likely to be tried in the future, although this is offset by a “fishing out” effect where the most valuable combinations of elements get used up. Elements that are embedded in a “good” network with lots of other pairs that have yielded useful ideas are also more likely to be used.

I investigate empirically these predictions using a large dataset on US utility patents: all 8.3 million utility patents granted between 1836 and 2012. I interpret a patent as a combination of elements that proved to be a useful idea. As a proxy for the elements that are combined to form ideas, I use the technology subclasses assigned to patents by the US Patent and Trademark Office (USPTO). Specifically, I use the hierarchical structure of the USPTO’s classification system in a novel way to aggregate more than 150,000 technology subclasses into 17,000 groups of comparable specificity. From this analysis, I find that the probability a pair of elements will be combined in any given year is (1) increasing in the number of past combinations, but at a decreasing rate, (2) decreasing over time,
and (3) increasing when both elements in the pair are also used with many other elements. These predictions are consistent with my model.

The same work also predicts that patenting activity is positively correlated with changes in researcher knowledge about the connections between elements, and negatively correlated with time. I investigate empirically this prediction using the same set of data on US patents, combined with the USPTO's technology classification system. I find that the growth rate of patents in each of 429 technology classes is falling over time, but that increases can be forecast from positive changes in my proxy for researcher information 1-5 years earlier, even after controlling for numerous other factors.

Going forward, I am also developing tests of this model's predictions about the complexity of ideas (the number of elements combined in an idea) and basic versus applied research over time. Moreover, the empirical analysis has many potential extensions. For example, it might be used to develop a good estimator of technological opportunity for a firm, industry, or country, or it could be used to study spillovers of information.

The second strand of my research examines the impact of environmental policy choice on innovation, when research is characterized by unobservable (to the policy-maker) variance in technological opportunity. I assume there exist two types of energy, clean and dirty, that are perfect substitutes except for their production costs and a negative externality from production using the dirty technology. Innovators are expected profit maximizers, and their decision to expend resources on R&D depends on technological opportunity, as well as the policy of the government.

We show that the policy-makers decision to use quota or price based incentives to spur innovation is not inconsequential. Price based incentives such as a carbon tax are characterized by disperse outcomes, with a lot of R&D resources expended when technological opportunity is high, and very low amounts when technological opportunity is low. Quotas, in contrast, lead to a more consistent level of R&D spending across differences in technological opportunity. Thus, price-based systems are more likely to deliver great technological advances or none at all, while quotas are more likely to deliver consistent incremental gains. Moreover, we also find a simple carbon tax is likely to outperform any quota in expected welfare terms.