Tom Hertel's influence and its lessons about academic inquiry

Russell Hillberry
Purdue University

David Hummels
NBER, Purdue University

October 2017

Abstract: Fields of academic inquiry differ in their preferred forms of output, in the ways in which knowledge is accumulated and stored, and so in the ways that academic influence is measured. We compare Tom Hertel's research record to other international economists of his generation in order to illustrate the unique breadth and influence of his work, and of the GTAP project broadly. We then provide an analytical framework that helps explain the evolution of the field of international economics from a tool-use standpoint. This framework helps us to assess the academic productivity gains from creating the GTAP model and consortium. It also provides a possible answer to a significant puzzle: why is GTAP increasingly influential in the physical and biological sciences, but less so within the international economics community?

Acknowledgements: We would like to thank Masha Brussevich for exceptional research assistance.

I. Introduction

It is a pleasure for us to be part of this project to honor Dr Tom Hertel. Tom is an outstanding economist in his own right, but his contributions cannot simply be judged by the standard metrics. They are much broader and important than that.

GTAP is a huge accomplishment. Many others have been involved, but the central contributor to the project is obviously Tom. He understood the need for the various GTAP products, assembled the team, and built a network of policy analysts around the world. We both have seen his years of patient attention to data issues, coding challenges, tractable representations of complex policies, and implications of functional form. But perhaps most incredible is his construction of a robust institution with a sustainable future that is assured by financial and technical support from leading policymaking bodies around the world. An achievement of this magnitude and scope is something that most academics would not even dare to attempt.

In this paper we offer a mix of theory and data to provide some perspective on these achievements. Our contributions are the following.

First, we undertake a citation analysis of Tom's academic work. In order to contrast the form of his impact with more standard research careers in economics we compare his citation record to three other leading trade economists. There are a number of notable differences, but most significant perhaps is the breadth of the impact across disciplines. Also notable is the late career shift into publications in the physical sciences, a shift that has substantially raised the profile of his publications as measured by impact factors.

In order to provide some context on the contributions of GTAP, we develop a simple theory of academic output and adduce relevant evidence to support its predictions. We describe a production function that combines multiple tasks – model building, solution algorithms, data collection, econometric estimation, and so on -- to generate results. Over time the substitutability of these tasks drops, making "single tool" papers decreasingly palatable in the eyes of the profession. For certain subjects such as trade policy analysis, production takes the "o-ring" formulation, in which all tasks are necessary. We provide analytics to show how the introduction of a public resource like GTAP significantly increases research productivity, particularly in applications where researchers have strong comparative disadvantage in necessary tasks.

The theory embeds a puzzle – why don't more academic researchers in the international trade field employ GTAP? And, more to the point, why don't authors writing in the "quantitative trade literature" recognize or acknowledge that they are replicating decades old developments in the Computable General Equilibrium literature? Our explanation is that these researchers are maximizing a different objective function, one that emphasizes both output and solo value added of the researcher. This yields a division of labor between GTAP and own-production that optimizes this augmented objective function, but is socially sub-optimal in terms of generating new findings.

II. Citation analysis

We begin by conducting a descriptive analysis of the publication and citation record of Professor Hertel, and to compare it with other prominent international trade economists as a point of reference. Our reference points are Professors Kyle Bagwell (Stanford), Jeffrey Bergstrand (Notre Dame) and James Tybout (Penn State). We chose these authors as comparisons because their careers are approximately the same length as Professor Hertel, and they have roughly similar numbers of citations in Web of Science. Moreover, each represents a markedly different methodological approach to international trade research: Bagwell writing theoretical papers, Bergstrand reduced form econometrics, Tybout structural econometrics, and Hertel Computable General Equilibrium approaches. Despite all working in the international trade field and having similar overall influence as measured by citations, Hertel's academic profile is qualitatively different from the rest of the group, and we wish to highlight these differences.

Specifically, Hertel's research has impact across a much broader range of academic fields. In the last decade in particular, Hertel's career output features a much closer connection – both inspired by and inspiring -- physical and biological sciences than have the other authors. This may be obvious to readers who have followed his career closely, but it is useful to quantify these differences in a transparent way. Our quantitative measures will include counts of publications, citations, and scientific fields in which each author is published and cited. We will also use the most recent impact factors from Web of Science as weights in order to normalize impact across disciplines.

Publications and Citation Counts

We take some data from Google Scholar, but most of the data we use in the analysis is taken from Web of Science™ (WoS). The data therefore only represent citations in WoS journals of publications in WoS journals. The advantage of this limitation is that it gives us a well-defined universe, with well-defined fields and a common approach to calculating impact factors. The approach is also limiting because some journals are not included in WoS, but benefits of a well-defined universe are worth the tradeoff.

To get a sense of the scale of this issue, we first report total numbers of citations in Google Scholar for each author and for Web of Science in Table 1. We also report the h index from Google Scholar.[1] All authors have far more cites in Google Scholar than in the Web of Science. The comparison authors have relatively consistent ratios of the two citation sources. Web of Science citations make up a far smaller share of Professor Hertel's Google Scholar citations. We do not find this particularly surprising, given Hertel's significant contributions to the policy analysis literature. Specifically, many policy analyses are not published in scientific journals and are omitted from WoS, but are found on the web and appear in GS. This also explains Hertel's significantly higher h-index. Since our primary purpose in this section is to describe Hertel's academic contributions, we focus next on the Web of Science publications and citations. We leave to the reader a consideration of whether analyses directly informing policymaking have greater/lesser impact, broadly defined, than those informing the academic world.

Table 1. Google Scholar citations, Web of Science citations and Google Scholar h-index

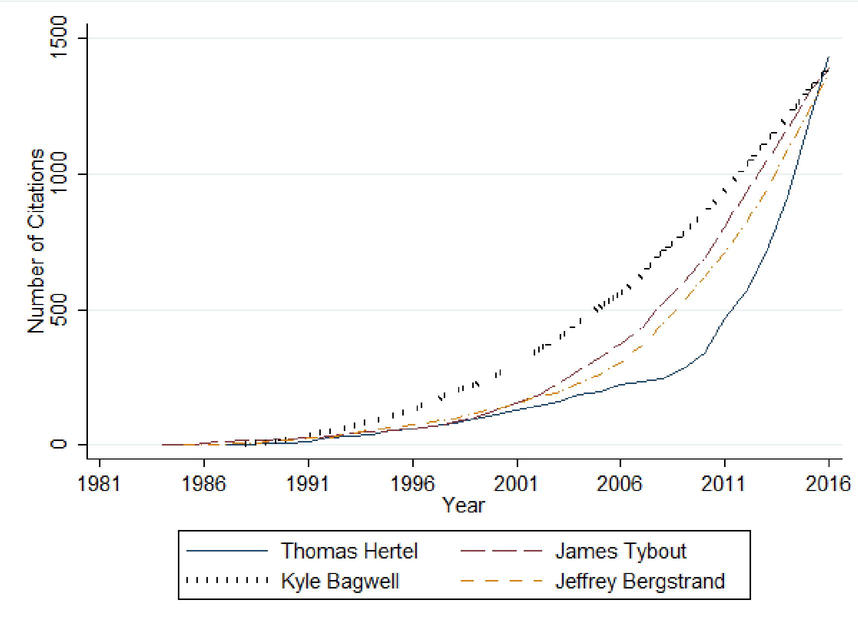
Author	**Google Scholar Citations**	**Web of Science Citations**	**h-index**
Thomas Hertel	20,098	1,454	70
Kyle Bagwell	11,167	1,395	48
Jeffrey Bergstrand	11,770	1,394	30
James Tybout	13,484	1,370	38

Google Scholar date of download: February 19, 2017
WoS date of download: March 7, 2017

It can be useful to understand how our four international economists arrived at these totals, so in Figure 1 we illustrate the citation histories over time. After significantly lagging our comparison group for roughly two decades, Hertel's citation count begins to accelerate around 2010, surpassing them in the past year.

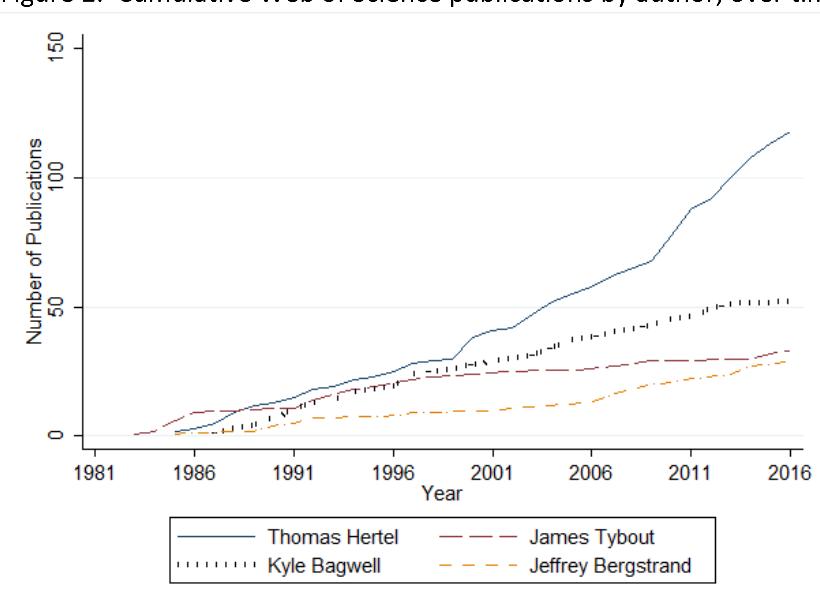
[1] The h-index is defined as the number of an author's publications with at least h citations.

Figure 1. Cumulative Web of Science citations by author, over time



The authors have different styles of work. One indication of these differences is visible in simple counts of publications as shown in Figure 2. While consistently higher than the comparison group through the first 15 years, Hertel's work undergoes an inflection point and accelerating publications from 2000 onward. While GTAP model and database had been operative prior to 2000, it is perhaps around this time that Hertel was able to begin really leveraging the project for his own work.

Figure 2. Cumulative Web of Science publications by author, over time



Breadth: Fields of Inquiry

A distinct feature of Tom's work is the extent to which it has reached a broad range of academic literatures. To indicate breadth of publications and citations we count the number of fields of inquiry in which each author is published/cited. To do this calculation we need something akin to a JEL code in Economics. Web of Science assigns journals to fields of inquiry, and we use these fields as our unit of reference.[2] Many journals are mapped to primary and secondary fields (and sometimes more).[3] We assign the field for each journal as the primary field as defined by WoS.

Although there is no doubt room for debate about the scope of specific fields of inquiry, we accept the WoS classification as definitive.[4] In order to provide the reader with some context about the breadth of fields included in the WoS designations, Table 2 lists the 30 most commonly cited fields for Hertel, also reporting the number of his WoS publications in each field, if any.

This appears to be an unusually broad reach for an international economist. How does it look in contrast to our comparison group? In figure 4 we report the cumulative number of distinct academic fields of publication, comparing Hertel to an average of the other 3. From 1990-2010 Hertel published in roughly twice the number of fields of the comparison set, and the gap accelerates in 2010.

A third point of distinctiveness is which fields use and cite these authors in their work. We use our mapping of journals to fields to assess which fields are citing each author, and display cumulative citation fields in Figure 5. Citations to Hertel's work appear in papers appearing in journals from a broader set of fields. The recent increase in Hertel's breadth (in terms of fields and citations) came with the shift into physical and biological sciences. Some important features of GTAP played a role here, a point we will explore more fully in the analytical model provided in the next section.

---

[2] Some journals, notably general interest scientific journals such as *Science* or *Nature*, are mapped by WoS to a field labelled "science, multidisciplinary" or "agriculture, multidisciplinary." In order not to overstate the breadth of an author's influence, in these cases we assign any such journal to a field that is already included in the author's field list. It does not matter for our calculations if a journal like *Science* is mapped to Biology, Chemistry, or any other field in the sciences.

[3] It is common for example that journals with a primary field designation in Agricultural Economics will be assigned a secondary field designation of Economics. In this example, the journal would be mapped to Agricultural Economics.

[4] We were not able to find documentation of the process for drawing field boundaries.

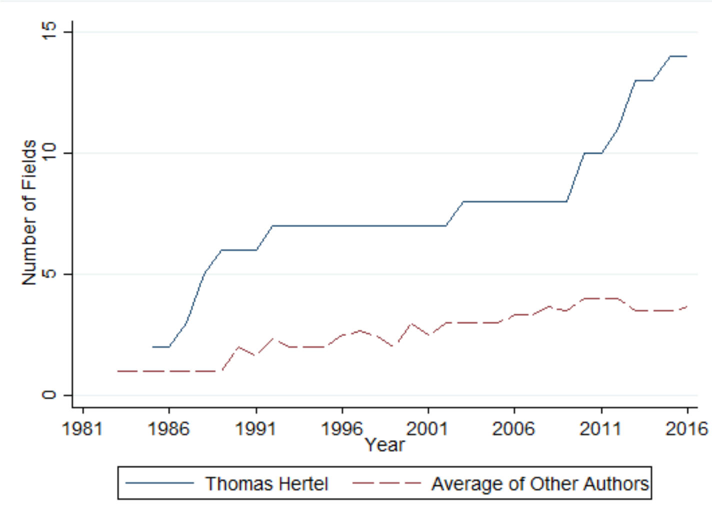
Table 2. Publications and citations by field for Thomas Hertel, top 30 fields by citations.

Field	Number of Publications	Number of Citations
Environmental studies	15	309
Economics	69	287
Agricultural economics & policy	5	130
Energy & fuels		73
Geography		62
Food science & technology	4	60
Multidisciplinary sciences	3	58
Planning & development	6	49
Meteorology & atmospheric sciences	1	49
Green & sustainable science & technology	1	39
Agriculture, multidisciplinary		39
Engineering		35
Agronomy		26
International relations	5	23
Ecology	3	21
Forestry		21
Biodiversity conservation		15
Geosciences, multidisciplinary		15
Biology	1	12
Management		9
Area studies		8
Biotechnology & applied microbiology		8
Chemistry		7
Agricultural engineering	2	6
Law		6
Plant sciences		6
Business & finance	2	5
Public, environmental & occupational health		5
Remote sensing		5
Soil science		5

Note: In subsequent calculations journals mapped to multidisciplinary fields are mapped to individual fields already included in the author's field list.

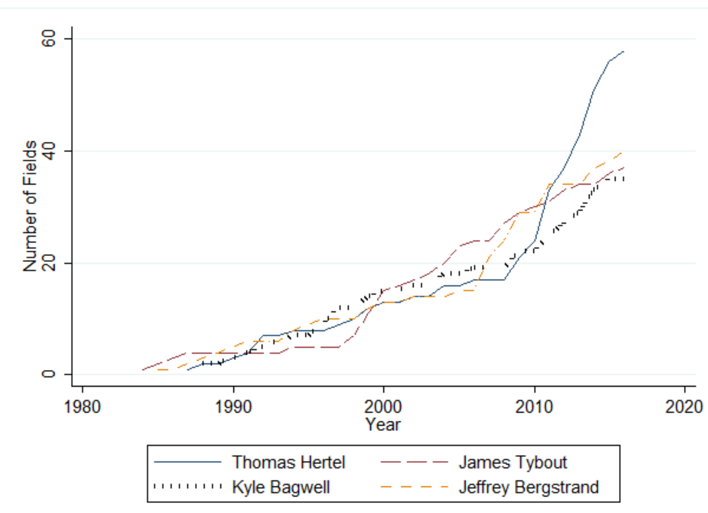
.

Figure 4. Number of Fields of publication, by author, over time.



Note: Other authors include James Tybout, Kyle Bagwell, and Jeffrey Bergstrand.

Figure 5. Number of Fields of citations, by author, over time.



Impact factors

Finally, we can measure impact not only with citations but by looking at the impact factors of the publications in which research appears. We use Web of Science impact factors that report, for each journal, the average number of times that an article published in the previous two years had been cited in those two years. The distribution of these statistics is skewed, both within fields and across them. To illustrate the quantitative scope of these statistics we report in Tables 3 all of Hertel's publications with impact factors greater than 3.75, a threshold that includes all of the "top 5" journals in Economics.

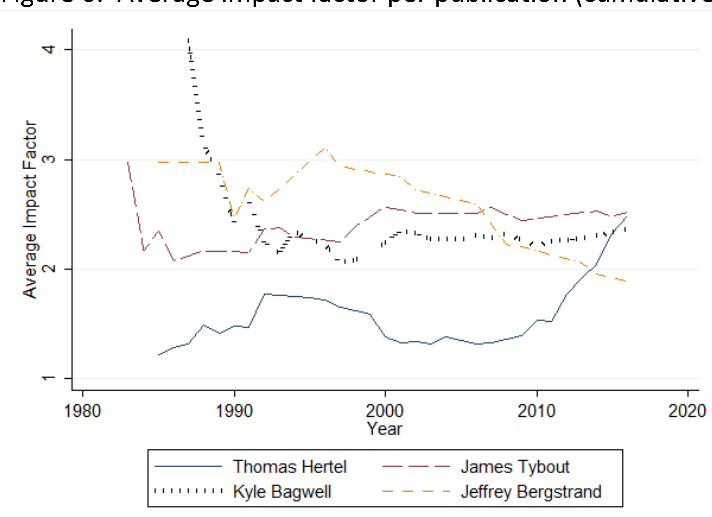
Table 3. Tom Hertel publications with high impact factors

Journal Title	**Journal Impact Factor**
ENVIRONMENTAL RESEARCH LETTERS (2009, 2013a, 2013b, 2013c, 2013d, 2014)	4.134
BIOSCIENCE (2010)	4.294
CURRENT OPINION IN ENVIRONMENTAL SUSTAINABILITY (2013)	4.658
ECONOMIC SYSTEMS RESEARCH (1992, 2004, 2016)	5.306
GLOBAL ENV. CHANGE-HUMAN AND POLICY DIMENSIONS (2010, 2011, 2014)	5.679
PROCEEDINGS OF THE NATL. ACADEMY OF SCIENCES OF THE USA (2012, 2014)	9.423
NATURE CLIMATE CHANGE (2012, 2016)	17.184
SCIENCE (2015)	34.661

For reference: Journal of Political Economy (3.75), American Economic Review (3.833), Econometrica (4.053) Review of Economic Studies (4.077), Journal of Economic Perspectives (5.012), Quarterly Journal of Economics (5.538), Journal of Economic Literature (6.614).

We report a comparison of the average cumulative impact factor in Figure 6. For much of this period, Hertel's research garnered significantly lower impact per publication. Figure 6 shows a marked shift, again beginning in 2010 with the transition into papers relevant to the physical and biological sciences. Average impact factors per publication catch up to the comparison authors by the end of the period. While it is beyond the scope of our study, it would be intriguing to learn whether there are other examples of prominent authors who have seen such a dramatic increase in impact factors 25 years into a research career.

Figure 6. Average impact factor per publication (cumulative)



In summary, Hertel's work is significantly broader than our comparison set authors, published in and cited by journals from a larger set of fields. This shift became most pronounced around 2010, at which point Hertel's citations and impact factors grew dramatically. This corresponds to a shift in his work to directly inform not just trade policy but the impact of the global economy on issues of import to the physical and biological sciences. We explore some of the reasons for this profound shift, and why Hertel was uniquely able to accomplish it, in the next section.

III. Some Simple Analytics of GTAP and International Trade Research

In this section of the paper we provide some simple analytics of GTAP. Our goal is to understand the evolution of the relevant literatures, as well as GTAP itself, to evaluate the social returns to GTAP, and finally to understand a significant puzzle. Despite the fact that GTAP provides a valuable public capital stock, many academic researchers investigating research questions wholly within the GTAP domain eschew its use. While this historically could be traced to skepticism about the CGE "black box", more recent work consists of creating entirely new black boxes that do many of the things GTAP can do. We seek to explain why. Naturally, we will do this using standard trade models!

III.A Stylized Facts

We need a foil for our argument so we will contrast the GTAP community with the NBER trade community, more specifically, the International Trade and Investment working group. Let's start with a few "facts" that we will buttress with qualitative arguments below.

1. **The scope of GTAP has grown considerably over time.** In Figure 7 we report growth in the number of contributors and consortium members over time. Figure 8 illustrates the growth in search results in Google Scholar for select keywords.

Figure 7. Number of active GTAP network and consortium members

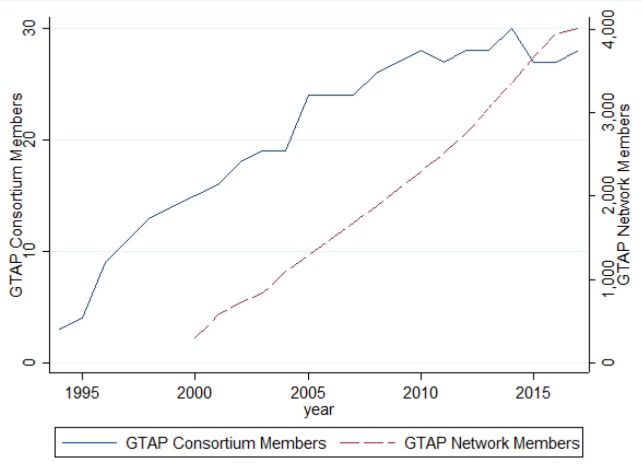
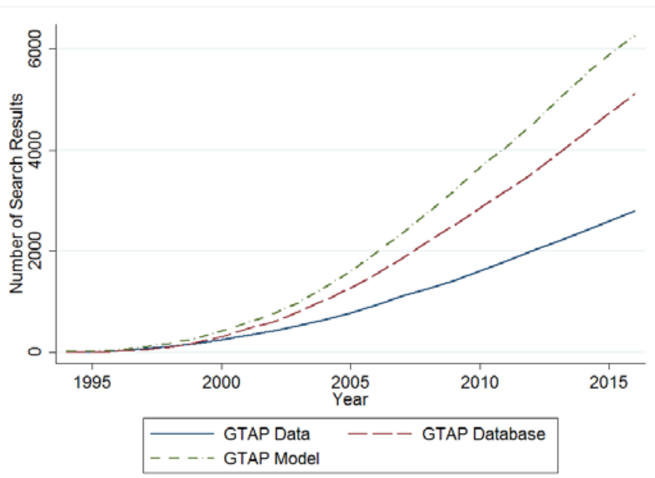
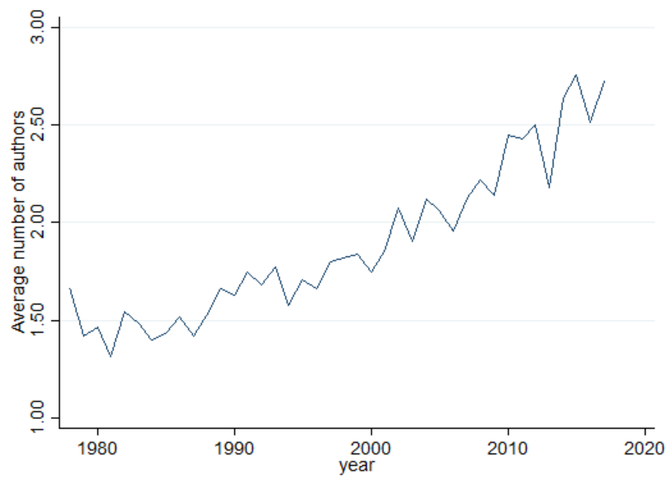


Figure 8. Number of Google Scholar search results for GTAP key words



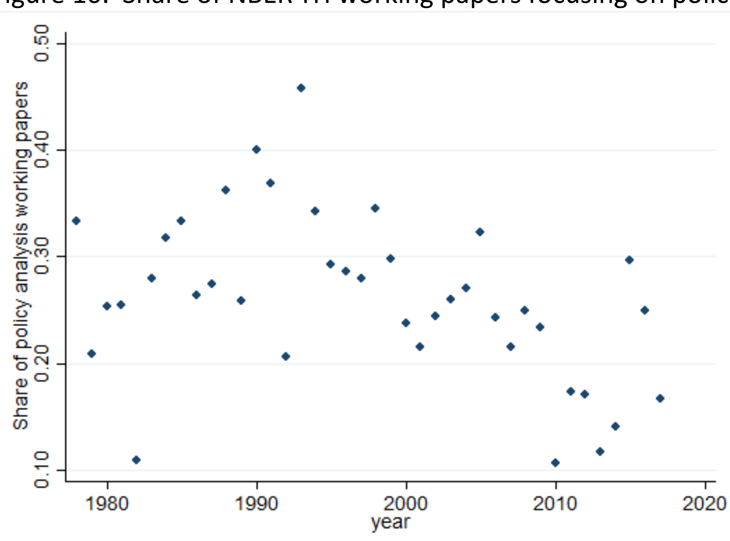
2. **Academic articles appearing in the NBER have become more complex**.   This can be seen in two key metrics. In Figure 9 we report the average number of coauthors on an NBER-ITI working paper and find that the number of coauthors has doubled in this period.

Figure 9. Average number of authors on NBER-ITI working papers



3. **Policy analysis is increasingly the province of CGE modelers, not the NBER**. In Figure 10 we report the number of policy analysis papers appearing in the NBER-ITI working paper series, both in total and as a percentage of all working papers in a given year.   The fraction of policy-related papers had fallen by almost half over a period of 20 years until a recent rebound in 2015 and 2016.[5]

Figure 10. Share of NBER-ITI working papers focusing on policy analysis



[5] This point is not unknown to NBER-ITI. So concerned were some ITI members about the dearth of policy analysis that the Spring 2014 NBER meetings were moved from their usual locale to Washington DC to try and spur some interest in policy!

4. **Complex CGE-style quantitative trade models have become prevalent in NBER conferences and working papers.**
   The advent of heterogeneous firms models in the style of Meltz (2003) and Eaton-Kortum (2002) have enabled a significant expansion of quantitative modeling. Two illustrative examples are Caliendo and Parro (2015) and Eaton, Kortum, Nieman, Romalis (2016). Like CGE-style papers, these papers often generate both positive and normative analysis of general equilibrium phenomenon. They are also, like CGE-style papers, complex and daunting to the uninitiated.

5. **GTAP output, but not NBER output, has been increasingly embraced in the scientific community.** The comparative citation analysis in the preceding section speaks to this point clearly.

III.B. A Simple Model of Academic Production: Extensive Margins

How do we understand these facts? Let's start with a simple model. Let y be the production of new results. A result could be a theorem, or a finding from a regression, or an analysis of the likely outcome of a trade policy change. Below we will talk about the mapping between results and professional rewards. Results are produced by researcher i using a combination of N tasks aggregated within a CES structure.

(1) $y_i = \left(\sum_n q_{in}^{\theta}\right)^{1/\theta}$

where $\theta = \frac{\sigma-1}{\sigma}$ and q_{in} is the quantity of task n produced by researcher i. A task is a constituent part of generating a result. Tasks include things like idea generation, construction of theoretical models, deriving solutions to those models, data collection, data analysis of various sorts, constructing visual displays of quantitative information, writing, and presenting. Of course, these tasks might be more finely subdivided to reflect specific kinds of models, solutions, datasets, estimation techniques, and so on.

Researchers differ in their productivity over tasks φ_{in}, and choose a quantity of each task to produce by devoting some portion of their total labor endowment to it.

(2) $q_{in} = l_{in} * \varphi_{in}$ *subject to* $L_i = \sum_n l_{in}$

Think about an instance where there are two tasks – theory and empirics – and in terms of generating results they are perfect substitutes ($\sigma = \infty$). Then the researcher would choose to produce results using the one task in which she enjoys greater productivity. If the tasks are imperfect substitutes, then the researcher will choose a mix of tasks weighted toward high productivity, with a mix of tasks that follows:

$$(3) \quad \frac{q_{theory}}{q_{empirics}} = \left(\frac{\varphi_{theory}}{\varphi_{empirics}}\right)^{\sigma}$$

Having strong theory productivity, at high levels of sigma, would generate results that are primarily theoretical in nature with a light dash of empirics thrown in. Perhaps a casual anecdote, or a motivating stylized fact.

Judging from what was published in good journals and what appeared in NBER working paper series, the international trade field was primarily a theoretical discipline for much of its history. What empirics did emerge in the 1970s and 1980s and even into the 1990s tended to be atheoretical and reduced form. Many papers consisted of simple measurement, others involved estimating simple aggregate gravity relationships or regressing intra-industry trade indices on correlates conjured up without the benefit of formal modeling. Empirics that were theory based had two problems. One, the work suffered from a severe lack of clarity about how to map theoretical predictions generated from low-dimensional models into complex high-dimensional data. Put another way, a 2-good 2-factor 2-country model without trade frictions might yield very clear theoretical predictions, but what they had to say in a 17,000-good, 5-factor, 150-country world where trade frictions abound is wholly unclear. Two, in other cases, like papers estimating the factor content of trade, the mapping from theory to data was exceptionally clear. But implementing these tests would founder because of dubious data (e.g. assuming all countries used the US technology matrix).

Note that as long as σ remains very high and the various tasks that go into the results production function remain close substitutes, this specialized state of the world could remain stable. Theoreticians here. Empiricists there. Measurement types over in the corner.

As is well known, production functions of this sort generate increasing returns to scale in the number of tasks. Holding $q_i = q \;\; \forall n$,

(4) $y = N^{(\sigma-1)/\sigma} * q.$

At a point in time, we can think of researchers as optimizing over a fixed set of tasks N. How might we endogenize N? Suppose we generalize the objective function for a researcher to include not just results, but citations. Citations depend on novelty but also usefulness, and can be thought of as the "prices" or professional valuation of a result. Novelty means that the more results there are, especially results within a similar area of work, the harder it is to generate citations. In a language that should be familiar to CGE modelers, continuing to produce the same type of result pushes readers down the demand curve for results, lowering citations or "prices", and worsening the terms of trade between research topics. Lowering the terms of trade between research topics is an episodic event in the trade literature. Each new innovation in topics (strategic trade policy, economic geography, multinational firms, political economy, heterogeneous firm theory) is met with a wave of theoretical papers extending, extending, extending and eventually driving the returns to those extensions (citations) to zero.

In this case, it might be "profitable" for some individual or institution to invest some fixed costs to invent a new task category instead of producing more results with the existing technology. This new task could be new kinds of data, a new flavor of preferences or GE modeling (Ricardo to Heckscher-Ohlin to Krugman to Melitz back to Ricardo), new solution algorithms or new estimation techniques. If these tasks have significant marginal productivity (i.e. enabling researchers to create distinctive results) they will be adopted in the production function and garner citations.

However, the incentives to invest in creating new tasks is quite weak as long as σ is large. In this case, the returns to incorporating that task into a researcher's production function would be quite low, which can be seen by evaluating (4). With large σ, researchers are nearly indifferent between incorporating another task in production or increasing quantities of existing tasks. The point could be made even stronger, and the returns made even lower, were we to incorporate some kind of dynamic "learning by doing" element to the task-specific productivity for that researcher. That is, researchers might have initially low productivity in using a new task technology and improve over time. But with high σ values, one might never find it useful to incorporate the new task.

But suppose σ begins to fall. Why? Perhaps scholars (and editors and referees) begin to regard theory by itself, or data by itself, as insufficiently persuasive. Perhaps a "result" is devalued if it is result #1000 exploring the permutations and combinations of strategic trade policy models, or economic geography models. This is especially important if the thicket of theoretical possibilities is not trimmed by using empirical regularities to select a preferred model. The same is true in terms of empirics -- by the time you've read a few hundred gravity or intra-industry trade papers dreaming up correlations without theoretical backing, interest dims.

As σ falls, a few interesting things begin to happen. First, the near exclusive reliance on one task, as captured by equation (3) for high σ, begins to become a more diversified production setting. Theorists might motivate models in terms of new stylized facts; empiricists ground their regressions in a theoretical prediction.   For even lower values of σ, significant contributions require both theory and solid empirics and often, unique datasets.   Casual empiricism suggests that it is increasingly rare to see high performing rookies on the PhD job market with "single tool" papers.

Second, because there is a rise in the use of new tasks, the returns to innovating new tasks begins to rise.   Examples abound. In response to new facts about differences in firm productivity and participation in global markets, Melitz (2003) developed a coherent framework for incorporating firm heterogeneity into a general equilibrium setting. Eaton and Kortum (2002) extended a Ricardian trade model to include many countries and trade costs and showed a tractable method for calibrating that model to higher dimensional data. Both these models generate aggregate gravity relationships but with zeros predominating in a full importer-exporter-product matrix. The PPML estimator proposed by Santos Silva and Tenreyro (2006) respond by providing a tractable method of incorporating zeroes in gravity estimation. These contributions have already generated 2165, 667, and 449 cites, respectively, in WoS, with many more to come. Each of these are results in their own right, but are far more important as tools that expand the economist's task set.

We hypothesize that demands for trade policy analysis play an important role in driving down market perceptions of σ.   If policymakers were content with theoretical predictions about gains from trade drawn from 2x2x2 analytical models, there would be little need to go further.   But policymakers want two things in particular that drive the need for task innovation.   They want specificity and they want quantification. Specificity means that aggregate statements (e.g. GDP rises by 0.1%) are less interesting than statements about particular sectors or particular regions or particular factors of production.

For example, author Hillberry was employed at the U.S. International Trade Commission in 2003 when the implications of removing the steel safeguard tariffs were explored. In the ITC analysis of that policy, aggregate gains from trade were not a central question, rather the key issues were the likely impacts of the tariffs on subsectors of the steel industry and on industries that were heavy users of steel (e.g. automobiles). The policy scenario was also complex, with variation in the level of the tariff over sectors as well as the exclusion of some import sources from the imposition of the tariff. Because safeguard tariffs are temporary, the policy was assessed in a framework that assumed limited adjustment of capital. Specificity also means that institutional context becomes relevant in a way that we are happy to gloss over when making aggregated statements.

Quantification means that signing derivatives isn't enough, particularly in cases where the theoretical effect is not clearly signed. The clearest example is tariff liberalization for a large country, where terms of trade losses may swamp efficiency gains from trade.

Greater specificity and quantification necessarily engages a greater number of tasks. Not just theory, but theory that can flexibly aggregate over many primary and material inputs and over many output sectors. Hat algebra for characterizing changes in equilibria in a concise way. And data! Lots and lots of data! Not just input-output tables, but IO tables specific to countries being analyzed. Accurate data on protection and shipping costs. Behavioral response parameters drawn from properly identified microeconometric estimation rather than being drawn from a hat. An understanding of the nuances of particular forms of trade liberalization more complex than reducing tariffs.

The authors of this paper have been part of this task generation process for CGE models from the early days. This includes data generation, parameter estimation, and incorporating novel modeling features.

Author Hummels compiled a large portion of the data behind the 1990 version of the Michigan Model by hand from a variety of UN and OECD data yearbooks. Heroic assumptions were made! Don't have the data on the number of Belgian textile firms because of data suppression? Well, assume the ratio of firms to output is the same in Belgium as in France, and suddenly a data hole is plugged. Multiply that sub-task 10,000-fold over the course of a year, and that is how one lonely grad student constructs the world economy (or a tenuous data representation of it) from scratch. This sort of thing is what happens when individual research teams, relying on callow grad students, try to build complex data sets. Consortium approaches are vastly superior. Data compilation is put in the hands of people who actually know the data they are working with, multiple teams working with the same data are more likely to uncover inconsistencies or find ways to improve coverage. The result is that consortium approaches are likely to enjoy vast improvements in productivity and quality control relative to isolated teams.

Much less arbitrary than the database work for the Michigan Model was the process of generating the CES elasticities of substitution now used in GTAP. It is well known in the literature that these parameters are central to both positive and normative evaluation of tariff liberalization in CGE modeling. Author Hummels had written a paper showing how to use bilateral variation in tariff and shipping costs to identify slopes of import demand curves. Hertel learned about this paper, and encouraged a rewrite focused on employing this technique to estimate substitution parameters for GTAP. Incorporation of proper microeconometric technique, and novel data sources, significantly improved GTAP parameter files. The estimates have been widely used outside the GTAP community, and Hertel, et al (2007) is the third most highly cited on Hertel's Google Scholar page.

Perhaps the greatest strength of the consortium is this ability to identify useful innovations and assimilate them, Borg-like, into the GTAP collective. Having a ready user base raises the returns on these innovations. For example, Meliz-style heterogeneous firm models have transformed theory and empirics outside of the CGE world. Author Hillberry co-authored a paper demonstrating methods for incorporating the Melitz technology into a Computable General Equilibrium framework like that in GTAP (see Balistreri *et al.*, 2011). Apart from using it to modify standard CGE questions related to trade and welfare responses, this task innovation expands the feasible set of results. Structural estimation techniques allow a recovery of the structural parameters from the Melitz theory that best fit the GTAP data.

III.C. When to Outsource Work to GTAP: Intensive Margins

Suppose an institution, let's call it "GTAP" gets into the business of not only innovating new tasks for the results production function, but providing these tasks directly to researchers. If a task requires the use of input-output tables from 100 countries employing consistent aggregation, a researcher can either construct those IO tables herself, or outsource the work to GTAP. Employing GTAP tools still requires some labor effort, but for many tasks it will require less effort than doing the work in autarky.

We are interested in two questions. One, how does the use of a GTAP technology improve results productivity? Two, how many tasks should a researcher do themselves? Part of the goal is to resolve a puzzling finding. GTAP is a research public good that enables researchers to engage in complicated multi-task research programs. It has been widely adopted in the policy analysis world. However, it is seldom used in the academic community whose scholarly work focuses on international trade. Why?

In order to illustrate the point analytically, we move to a pure Leontief representation of the complementary between tasks, that is, we evaluate equation (1) as $\sigma \rightarrow 0$. Pure complementarity allows a representation with a continuum of tasks, which allows in turn the use of integral calculus methods to determine the range of tasks undertaken with GTAP products.

Let $\varphi_i(n)$ be the productivity of researcher i at task n. For the purposes of subsequent analysis we will order the productivities of tasks $\in [1, N]$, and apply a functional form for $\varphi_i(n)$, namely $\varphi_i(n) = \varphi_i n^2$, a form that allows a consistent ranking of authors i The labor required to produce task n is therefore $\frac{1}{\varphi_i n^2}$and the labor required to produce all tasks in the analysis is $\sum_{n=1}^{N} \frac{1}{\varphi_i n^2}$. For what follows, a representation with a continuum of tasks is more tractable than the summation in equation (1) , so we write the labor required to produce an entire unit of analysis as $\int_{n=1}^{N} \frac{1}{\varphi_i n^2} dn$.

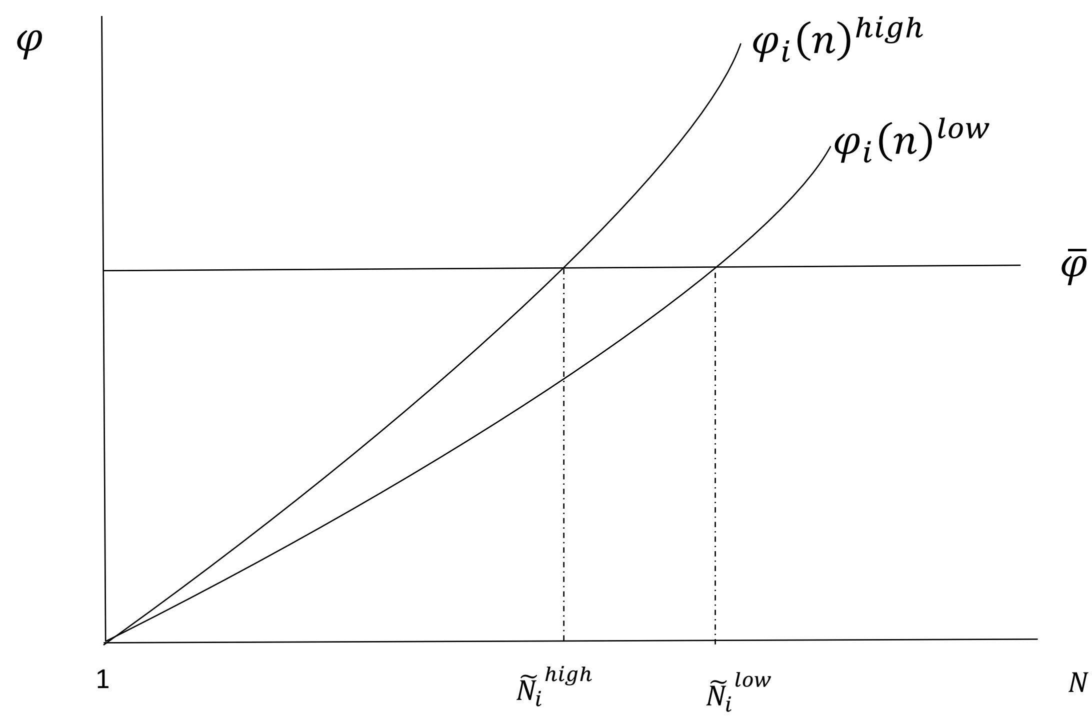
Assuming a fixed stock of labor for each researcher L_i, we can solve for the equilibrium number of papers produced by researcher i as a function of the total number of tasks required N, and researcher productivity φ_i:

(5) $$y_i = \frac{L}{\int_{n=1}^{N} \frac{1}{\varphi_i n^2} dn} = \frac{L\varphi_i N}{N-1}$$

This is our benchmark prediction of researcher output in the absence of GTAP. In order to incorporate GTAP in the model we assume that there are a range of tasks for which GTAP raises researcher productivity. The tasks may include the assembly of a consistent data set with input-output relationships, the assembly of tariff data, computational solution of a GE model via hat calculus, etc. For simplicity, assume each task conducted using GTAP is done with productivity level $\bar{\varphi}$, regardless of the researcher's individual productivity. For each researcher there is a cutoff value of N, $\widetilde{N}_i$, that determines which tasks are undertaken with GTAP and which tasks are undertaken without GTAP.

We illustrate that in Figure 11.

Figure 11. Researcher productivity and the measure of tasks undertaken with GTAP.



Given the ordering of task productivity for researcher i, for $\varphi_i < \bar{\varphi}$ there will be a range of tasks (1 to $\tilde{N}_i$) that the author chooses to do with GTAP, and a range of tasks ($\tilde{N}_i$ to N) that are done without GTAP. Applying our functional form for researcher i's task probability, we can solve analytically for $\tilde{N}_i$:

(6) $\tilde{N}_i = \left(\frac{\bar{\varphi}}{\varphi_i}\right)^{0.5}$

The higher is the researcher's individual productivity, the fewer tasks she will undertake using GTAP. Moreover, the cross partial of this value with respect to $\bar{\varphi}$ and for φ_i is negative, indicating that an increase in productivity by GTAP generates a smaller increase in the range of tasks undertaken by the high productivity researcher than the low-productivity researcher. This insight is apparent in the figure.

We now solve for the number of researcher i analytical outputs when the GTAP technology is available to them. Our primary interest is in the interior solution:

$$(7)\quad y_i = \frac{L}{\int_{n=1}^{\tilde{N}_i} \frac{1}{\bar{\varphi}} dn + \int_{n=\tilde{N}_i}^{N} \frac{1}{\varphi_i n^2} dn} = \frac{L}{\frac{\tilde{N}_i - 1}{\bar{\varphi}} + \frac{N - \tilde{N}_i}{N \tilde{N}_i \varphi_i}} = \frac{L N \tilde{N} \varphi_i \bar{\varphi}}{N \tilde{N} \varphi_i (\tilde{N}_i - 1) + \bar{\varphi}(N - \tilde{N}_i)}$$

One can check our earlier solution for $\tilde{N}_i$ by choosing $\tilde{N}_i$ to maximize y_i. After some moderately painful algebra one finds $\tilde{N}_i = \left(\frac{\bar{\varphi}}{\varphi_i}\right)^{0.5}$, as expected.

On the effects of changes in GTAP Productivity

One useful application of the model is to understand how an improvement in GTAP affects researcher productivity. In principle an improvement in GTAP might operate either through extensive or intensive margins of GTAP activity.[6] We consider an intensive margin, an improvement in the productivity of tasks undertaken with GTAP. An improvement of GTAP productivity has two effects: an inframarginal effect on productivity of tasks already undertaken with GTAP, and an extensive margin in terms of the number of tasks undertaken. The effects on the extensive task margin are immediate from the preceding equation. A ten percent improvement in task productivity generates a 5% increase in the measure tasks undertaken by researcher i. Infra-marginal effects of a productivity change on y_i are directly proportional to the share of tasks that are already undertaken using GTAP.

In order to combine these effects we substitute the solution for $\tilde{N}_i$ into the formulation for y. Applying the algebra, we are able to solve for y_i in terms of parameters:

$$(8)\quad y_i = \frac{L N \varphi_i^{0.5} \bar{\varphi}^{0.5}}{2N - N \varphi_i^{0.5} \bar{\varphi}^{-0.5} + \bar{\varphi}^{0.5} \varphi_i^{0.5}}.$$

[6] For example, the introduction of dynamic modeling capabilities into GTAP (Walmsley and Ianchovichina 2012) might be viewed as improvement along the extensive margin because it offers new capabilities. An intensive margin improvement might be, for example, an update of the database or the inclusion of more countries into the database.

The marginal effect of an increase in GTAP productivity is positive

(9) $\frac{\partial \mathrm{y_i}}{\partial \bar{\varphi}} = \frac{\mathrm{LN}\varphi_i^{0.5}\bar{\varphi}^{-0.5}}{\left(2\mathrm{N}-\mathrm{N}\varphi_i^{0.5}\bar{\varphi}^{-0.5}+\ \bar{\varphi}^{0.5}\varphi_i^{0.5}\right)^2} > 0.$

One can also evaluate the heterogeneous effects (across researchers) of a shock to $\bar{\varphi}$. The cross partial of researcher output with respect to and own productivity $\left(\frac{\partial^2 y_i}{\partial \bar{\varphi} \partial \varphi_i}\right)$ does not have a definite sign. On the one hand, a low productivity researcher conducts more activities with GTAP, and see larger increases in the share of activities conducted with GTAP when GTAP productivity is increased. On the other hand, labor savings for the more productive researcher are worth more in terms of total output.

Different approaches to valuing research output

An observable feature of the emerging quantitative trade literature is that it is relearning many of the lessons of the early CGE literature. Most obviously, the standard attribution of hat calculus solution methods to Dekle, Eaton and Kortum (2008) belies the adoption of these methods by the CGE literature long ago.[7] But there are other inefficiencies in the research process that could be addressed by using GTAP products, most especially in the collection and production of data. With some notable exceptions (Johnson and Noguera 2012, Caron Markusen and Fally 2014) there have been very few authors that have used GTAP products in high profile academic research. One key purpose of our model is to explain why.

Suppose that the academic market is not merely interested in obtaining research outputs but also interested in identifying highly productive researchers. The existence of a GTAP network that can be used to address many of the research questions reduces the cost of doing trade and welfare calculations. But authors using GTAP inputs in that exercise would dramatically complicate the effort of academic economics to evaluate researchers' multidimensional productivity.

In order to illustrate this in our model, suppose researchers face an academic utility function that rewards output, but also the contribution of the researcher to the relevant value added. Evaluating

[7] Hat calculus solution methods for computable general equilibrium date to Johansen (1960), and made their way into GTAP through the ORANI model in Australia. We thank Peter Dixon for providing us with a short history. We have also found it notable that some researchers in the quantitative trade literature seem even to be unfamiliar with the canonical application of hat calculus in the academic trade literature, Jones (1965).

value added is a general problem faced by any multi-coauthor team, or when separating the work of advisors from advisees on job market papers. But when it comes to separating value added provided by the researcher from that provided by GTAP consortium tools, authors reveal the division very clearly to the profession.

We parameterize this utility function as $U = \frac{N}{\tilde{N}} y$, which weights negatively use of labor saving tools such as GTAP. Substituting equation (7) into U returns

(9) $U_i = \frac{N}{\tilde{N}_i} y_i = \frac{LN^2 \bar{\varphi} \varphi_i}{\tilde{N}_i N \varphi_i (\tilde{N}_i - 1) + (N - \tilde{N}_i) \bar{\varphi}}$

We can find optimized values of $\tilde{N}_i$ by minimizing the denominator. The solution for the value of $\tilde{N}_i$ that optimizes U_i is

(10) $\tilde{N}_i = \frac{1}{2} + \frac{\bar{\varphi}}{2N\varphi_i}$

This assumes an interior solution, but one can also check whether the author will choose to use GTAP at all in this circumstance. Note that the lower limit for N is 1. The threshold for use of GTAP $(\tilde{N}_i \geq 1)$ will not be met so long as $N > \frac{\bar{\varphi}}{\varphi_i}$. Individuals with relatively high productivity will eschew GTAP altogether. While this reduces the quantity of research outputs that are produced, disincentives against use of GTAP (and presumably other similar tools) improves the field's ability to evaluate task productivity levels. Put another way, while raw productivity is more easily uncovered if authors eschew GTAP, there is a cost in terms of the progress of knowledge.

GTAP and the physical sciences

Evidence from Hertel's citation patterns above suggests that the physical and biological sciences have aggressively cited and used GTAP tools even while Economics shies away. In terms of the analysis above, the Sciences seem to be optimizing y_i rather than U_i. While speculative, we think that there may be a few explanations for this behavior.

One possibility is that the traditional sciences have a greater focus on replicability than does Economics. An important feature of a publicly validated data set, model, and solution algorithm is that replicability is feasible. It is plausible to imagine two GTAP users employing the tool to evaluate, say, how liberalizing sugar trade with Australia will affect land use, and arrive at the same set of conclusions. Or at least to definitively trace differences in conclusions to a specific, and replicable, choice of model or aggregation.

Somewhat ironically, an important argument for single tool Economics research was that it could be replicated. A piece of mathematical theory is, given assumptions, either right or wrong and can be independently verified with a little algebra. Simple econometrics using well-known data sets are similarly replicable, especially when authors provide data and associated code. In the authors' view, the profession's preference for replicability in single tool Economics generated a good bit of early distaste for and criticism of "black box" CGE work. And in the early days, this criticism was no doubt warranted! But that is a very different situation than the state of affairs now with a robust consortium of users validating data and model and solutions.

In contrast, the authors of this paper are not entirely confident that two authors attempting to use new style quantitative trade models to analyze the same policy experiment would wind up in the same place. Different data, different modeling choices, different solution algorithms, each tucked in their own tidy black box. (And don't get us started on the ability of a referee to verify whether these elegantly complicated constructions are "right".)

A related factor might be the routine use of laboratories to conduct research in the physical and biological sciences. Disciplines that have a lab tradition, and an emphasis on discrete contributions by large multi-author teams, might more naturally find the output of a consortium more credible than a tradition built on lone researchers.

Another possibility is suggested by recent survey evidence reported in Leslie et al (2015), which indicates that Economists view raw innate ability as more important for success than do practitioners of all the sciences except Physics. An academic culture that values brilliance might generate incentives to value U rather than Y.

IV. Conclusion

Many festschrift papers provide assessments of a celebrated author's insights in specific papers or the accumulated weight of their contributions to literatures they helped establish or shape. We set out to provide such an assessment for Tom Hertel's career contributions to academic discourse in Economics and beyond. But we wound up in such a different place that to evaluate his contributions in a paper or a literature seems like a category error. We think the appropriate way to evaluate Tom's contributions is to ask, as a profession, what are we trying to accomplish?

Likely many scholars would say that we are trying to seek a deeper understanding of the economy, and a more accurate assessment of how particular policies affect it. But what is the best way to get there? Is it generating more results (y in our formulation) or greater replicability in our empirical assessments? If so, it seems that the approach that Tom Hertel has taken with GTAP is an important step in this direction. It enables researchers to emphasize their comparative advantage, while outsourcing their relatively low productivity tasks to the consortium. It enables complete replicability of results, and iterative experimentation with various aspects of modeling to understand sensitivity of results to conditions. It is more like the way that knowledge is pursued in the sciences, and that no doubt contributes to why GTAP has been employed there.

Of course, generating more results that represent marginal iterations on the same theme may run into sharply diminishing returns, or rather, a depreciation in the terms of trade. Some might say that a deeper understanding of the economy comes from arbitrage between fields. Increasingly, the big challenges we face – climate change, energy -- require insights that cross disciplines. They involve understanding the physical environment in which we live, and the economic environment that conditions any potential change. Here too GTAP stands tall, because progress requires innovation in task space, and a willingness to acknowledge that a small contribution is still valuable in the larger process of scientific advancement.

All of this pushes in the direction of greater complexity of tasks and organization. In our simple analytics we described how increasing the task space, and outsourcing portions of that task space, yield substantial productivity gains. But all that comes at the cost of rising complexity, and one might reasonably question whether complexity necessarily adds insight. Consider this counterfactual. Suppose Paul Krugman had been writing his seminal articles today, and been "forced" by the tastes of the profession to calibrate and quantify his models of monopolistic competition or economic geography. Would we have learned more from that exercise? Would he have been more productive in generating insight?

We don't really know the answers to these questions. What we do know is that Tom Hertel has charted an interesting and meaningfully different career path. Different people will assess it differently. The physical and biological sciences, for example, are paying attention to it in a way that closely related areas of Economics has not. But assessing that career path has caused us to really start asking ourselves: what exactly is a contribution? And to conclude that, however one defines it, Tom Hertel's contributions are abundant indeed.

References

Ahmed, Syud, Noah Diffenbaugh, Thomas Hertel, David Lobell, Pedram Rowhani, Navin Ramankutty and Ana Rios. “Climate Volatility and Poverty Vulnerability in Tanzania.” *Global Environmental Change* 21.1 (2011), pp. 46-55.

Ahmed, Syud, Noah Diffenbaugh and Thomas Hertel. “Climate Volatility Deepens Poverty Vulnerability in Developing Countries.” *Environmental Research Letters* 4:034004 (2009).

Baldos, Uris and Thomas Hertel. “Looking back to move forward on model validation: Insights from a global model of agricultural land use.” *Environmental Research Letters* 8:034024 (2013a).

Balistreri, Edward J., Russell H. Hillberry, and Thomas F. Rutherford. "Structural estimation and solution of international trade models with heterogeneous firms." *Journal of international Economics* 83, no. 2 (2011): 95-108.

Caliendo, Lorenzo, and Fernando Parro. “Estimates of the Trade and Welfare Effects of NAFTA.” *The Review of Economic Studies* 82.1 (2015), pp. 1-44.

Caron, Justin, Thibault Fally, and James R. Markusen. "International trade puzzles: a solution linking production and preferences." *The Quarterly Journal of Economics* 129.3 (2014), pp. 1501-1552.

Dekle, Robert, Jonathan Eaton, and Samuel Kortum. "Global rebalancing with gravity: measuring the burden of adjustment." *IMF Economic Review* 55.3 (2008), pp. 511-540.

Diffenbaugh, Noah, Thomas Hertel, Martin Scherer, and Monika Verma. “Implications of Climate Volatility for Agricultural Commodity Markets under Alternative Energy Futures.” *Nature Climate Change* April 22 (2012), DOI: 10.1038/NCLIMATE1491.

Eaton, Jonathan, and Samuel Kortum. "Technology, geography, and trade." *Econometrica* 70.5 (2002), pp. 1741-1779.

Eaton, Jonathan, Samuel Kortum, Brent Neiman, and John Romalis. “Trade and Global Recession.” *The American Economic Review* 106.11 (2016), pp. 3401-3438.

Golub, Alla, Benjamin Henderson, Thomas Hertel, Pierre Gerber, Steven Rose and Brent Sohngen. “Global Climate Policy Impacts on Livestock, Land Use, Livelihoods and Food Security.” *Proceedings of the National Academy of Sciences* 110.52 (2012), pp. 20894-20899.

Hertel, Thomas. “Food Security Under Climate Change.” *Nature Climate Change* 6 (2016), pp. 10-13.

Hertel, Thomas, Navin Ramankutty and Uris Baldos. “Global market integration increases likelihood that a future African Green Revolution could increase crop land use and CO2 emissions,” *Proceedings of the National Academy of Sciences* 111.38 (2014), pp. 13799–13804.

Hertel, Thomas, Marshall Burke and David Lobell. “The Poverty Implications of Climate-Induced Crop Yield Changes by 2030.” *Global Environmental Change* 20.4 (2010), pp. 577-585.

Hertel, Thomas, Alla Golub, Andrew Jones, Michael O’Hare, Richard Plevin and Daniel Kammen. “Global Land Use and Greenhouse Gas Emissions Impacts of US Maize Ethanol: Estimating Market-Mediated Responses,” *BioScience* 60.3 (2010), pp. 223-231.

Hertel, Thomas, David Hummels, Maros Ivanic, and Roman Keeney. "How confident can we be of CGE-based assessments of Free Trade Agreements?" *Economic Modelling* 24.4 (2007), pp. 611-635.

Hussein, Zekarias, Thomas Hertel and Alla Golub. “Climate change mitigation policies and poverty in developing countries”, *Environmental Research Letters* 8: 035009 (2013b).

Johansen, Leif. *A multi-sectoral study of economic growth*. Vol. 82. Amsterdam: North-Holland, 1960.

Johnson, Robert C., and Guillermo Noguera. "Accounting for intermediates: Production sharing and trade in value added." *Journal of International Economics* 86.2 (2012), pp. 224-236.

Jones, Ronald W. "The structure of simple general equilibrium models." *Journal of Political Economy* 73.6 (1965), pp. 557-572.

Leslie, Sarah-Jane, Andrei Cimplan, Meredith Meyer, and Edward Freeland, “Expectations of brilliance underlie gender distributions across academic disciplines.” *Science* 347.1629 (2015), pp. 262-265.

Liu, Jianguo, Harold Mooney, Vanessa Hull, Steven Davis, Joanne Gaskell, Thomas Hertel, Jane Lubcheco, Karen Seto, Peter Gleick, Claire Kremen and Shuxin Li. “Systems Integration for Global Sustainability.” *Science* 347.6225 (2015).

Liu, Jing, Thomas Hertel, Farzad Taheripour, Tingju Zhu and Claudia Ringler. “International Trade Buffers the Impact of Future Irrigation Shortfalls.” *Global Environmental Change* 29 (2014), pp. 22-31.

Lobell, David, Uris Lantz Baldos and Thomas Hertel. “Climate Adaptation as Mitigation: The Case of Agricultural Investments.” *Environmental Research Letters* 8:015012 (2013c).

Melitz, Marc J. "The impact of trade on intra-industry reallocations and aggregate industry productivity." *Econometrica* 71.6 (2003), pp. 1695-1725.

Meyfroidt, Patrick, Eric Lambin, Karl-Heinz Erb and Thomas Hertel. “Globalization of land use: Distant drivers of land change and geographic displacement of land use.” *Current Opinion in Environmental Sustainability* 5.5 (2013), pp. 438-444.

Peters, Jeffrey and Thomas Hertel. “Matrix Balancing with Unknown Total Costs: Preserving Economic Relationships in the Electric Power Sector.” *Economic Systems Research* 28.1 (2016), pp. 1-20.

Reimer, Jeff and Thomas Hertel. “Estimation of International Demand Behavior for Use with Input-Output Based Data.” *Economic Systems Research* 16.4 (2004), pp. 347-366.

Silva, JMC Santos, and Silvana Tenreyro. "The log of gravity." *The Review of Economics and statistics* 88.4 (2006), pp. 641-658.

Steinbuks, Jevgenijs and Thomas Hertel. “Energy prices will play an important role in determining global land use in the twenty first century.” *Environmental Research Letters* 8:014014 (2013d).

Tsigas, Marinos, Thomas Hertel, and James K. Binkley. "Estimates of Systematic Reporting Biases in Trade Statistics." *Economic Systems Research* 4.4 (1992), pp. 297-310.

Verma, Monika, Thomas Hertel and Noah Diffenbaugh. “Market-oriented Ethanol and Corn-trade Policies can Reduce Climate-induced US Corn Price Volatility.” *Environmental Research Letters* 9:064028 (2014).

Walmsley, Terrie L., and Elena Ianchovichina. *Dynamic modeling and applications for global economic analysis*. Cambridge University Press, 2012.