Some Thoughts on Empirical Research

Li Gan

March 2008

I’ll start with a story in the high-school physics: the story about Isaac Newton’s Law of Universal Gravitation.

How Newton starts to think about the law of gravitation? The story was that he watched an apple falling down the earth, and started to wonder why the apple does fall down the earth. According to the story, the brilliance of Newton was he asked the question.

Unfortunately, how to ask a brilliant question cannot be taught in class. But fortunately, for most people, asking a brilliant idea is not necessary to be very successful.

The next step is how to transform the question (such as “why an apple falls down the earth”) into an abstract question (such as “why any object would fall down on the earth”), and then how to form a model to study this abstract question. Similarly, we cannot teach this part, neither. However, this part of research can be self-learned to some extent by enough experience: reading papers and doing research ourselves.

The part that can be taught is how to solve the model once you have one. That’s where most of your training is about.

Most empirical research actually comes after a model or a hypothesis has been constructed. It often answers a different question: How can the gravitation model be established as a theorem? To establish a model or a hypothesis to be true, several steps are necessary.

First, the model or hypothesis has to be able to explain the observed key relationships. This is absolutely the minimum requirement. So the Newton’s model has to be able to explain why apple falls on the earth but the moon does not.

Second, it has to be established that alternative hypotheses or models are wrong while only the concerned model or hypothesis is correct. For example, in terms of the gravitation theory, it has to be established that the alternative explanation on why apples fall down the earth is false. Suppose we just say an alternative hypothesis is that everything has to fall down on earth by some unrecognized force. Newton’s gravitation law actually says that not every object will fall down on the earth – it is related to the speed that the object travels. For example, it says moon would not fall down the earth, as we observed.

Third, also as a further validation of the model or the hypothesis, the model or the hypothesis has to have some predictive power. The model or the hypothesis can be
applied to predict phenomena or relationships that are previously not known. For example, people applied Newton’s theory to predict the existence of Neptune and Pluto.

The story was, Le Verrier used Newton’s laws of gravitation to predict the existence and the orbit of Neptune. He mailed his calculation by September 18, 1846, to Johann Galle or the Berlin Observatory. The letter arrived five days later and the planet was found the same evening. This completes the test of the Universal Laws of the Gravitation.

Testing alternative theories is by far the most important-type of empirical studies. Other types of empirical studies include finding a relationship to guide the theoretical investigation, and estimating the parameters of a model.

Given these discussions, the empirical studies in economics consist of four broadly defined categories.

(1) Testing alternative models.
   Almost all economic models imply some forms of causality. So testing alternative models involves testing causality. Specifically, there are two issues involved: first, deriving different causal relationships among different models or hypotheses; second, estimating which causal relationship is correct.

   To establish or to estimate a causal relationship, one has to hold all other (relevant) factors fixed. For example, to find if having another year of education causes an increase in monthly salary, one has to hold family background and personal characteristics constant.

   A large part of econometrics is about establishing causal relationships. Almost all regressions in the type of \( E(y|Z) = f(X|Z) \) assume a change in \( X \) causes a change in \( y \) where \( Z \) is the set of exogenous variables.

   I will use an example that relates to one of my current research projects. This example is not necessarily the best example available, but it is easier for me to talk about.

   Example: A positively correlated housing price and transaction volumes. Stein (1995) found that housing prices and transaction volumes are positively correlated empirically. A decrease of 10% in price lowers transaction volumes by 1.6 million units, which is about 40% of total transaction volumes.

   Alternative theories: the down payment model; the loss-aversion model; and the thin-thick model.

   It is possible that all models are correct, or any of them is correct.
The down-payment model: when the housing price is down, the equity from the house may not cover the down payment of the new house. So people who want to move cannot.

The loss-aversion hypothesis: marginal disutility from a loss is larger than the marginal utility from a gain. Sellers tend to hold their house in hope of offers higher than the original purchasing prices when market is down.

Thin-thick market hypothesis: when unemployment is up, a lower number of buyers is on the market, that instead causes a lower number of sellers on the market. The market is thinner. The match quality between buyers and sellers are lower, and hence it takes longer time to find a match.

They all tell the same story, however, the loss-aversion hypothesis and the down-payment model both only apply to a down market. The thin-thick market hypothesis applies to both down market and up market. This gives a test for alternative models.

(2) Using existing theories to predict previously unknown phenomena. This can be considered as a further validation of the existing theories.

The thin-thick hypothesis can only be used to study the transaction volume relationships, it can predict: (a) a larger housing will have a shorter time-to-sale, a smaller price dispersion, and a higher average sale price; (b) when unemployment rate goes up, the sale price decreases by a smaller percentage in a larger housing market.

Many papers that combine theory and empirical work should fall in this category. The requirement for this type of empirical research is quite high:

First, the theory should be motivated by a known empirical phenomenon, and can be used to explain this empirical phenomenon.

Second, the theory can be used to derive additional predictions that are previously unknown and that can be empirical verified.

Finally, the empirical work. If there are already hypotheses available to explain the empirical phenomenon, then it is often necessary to test which one is correct.

Newton example: Finding Neptune and Pluto

Often a paper that is combined with a theory and an empirical work is judged by the minimum, not the sum.
For both (1) and (2), since they are heavily involved with theories, often the causality as implied by theories are crucial. A huge part of econometrics then tries to deal with the causality.

(3) Obtaining relationships between variables.

a. Obtaining the precise magnitude of the existing relationship.
   Example 1: the return to education. One year increase in education can lead how much increase in wage.
   Example 2: the elasticity of labor supply (with respect to wages). A 1% increase in wage can lead to how much increase in labor supply.

   Obviously, the magnitude has to be very interesting and very crucial.
   Examples: Time discount rate
   Relative risk aversion rate
   Bequest motivation
   ...

b. Obtaining statistical relationships between two variables. The bar of this type of relationship is high: the obtained-relationship has to be rather surprising.

   Example 1: Explaining international trade
   (i) Traditional theory of comparative advantages: each country would export the good that it could produce at lower relative cost in autarky. The Heckscher-Ohlin model, based on relative differences of primary factor endowments, came to dominate textbooks as well as research papers. Here each country had comparative advantage in the good that used relatively more intensively its relatively more abundant factor.

   (ii) However, a series of empirical papers since 1960s show that the fastest growing component of trade was between industrial countries with very similar factor endowments. The European Common Market brought together countries that were not complementary in their factor endowments. Much of this trade expansion seemed to occur with relatively little distributive conflict within each country. Finally, in many emerging industries, one could not point to a clear comparative advantage for any country. Many patterns of production and trade seemed matters of chance; in fact there was a lot of two-way trade in very similar products such as automobiles.
These empirical findings are important because they pose a major challenge to the existing theory of Heckscher-Ohlin model. It directly leads to a series of new theories of international trade, called New International Trade.

Example 2: Davis and Haltiwagner (1995): job destruction and job creations. They argue that a lot of jobs are created and destroyed at the plant level.

Example 3: The relationships between the income and consumption. According the Keynesian, the percentage that people spend on consumption has to be low as income increases. However, Kuznet found that this percentage did not decrease or increase over a long period of time.

Example 4: Deaton and Paxson (1998)
Food consumption in total consumption is negatively correlated with per capita income. Deaton and Paxson claims this is paradoxical, since food consumption should be positively correlated with per capita income.

The obtained relationships have to be very important. The importance often is judged by theory.

(4) Structural estimates of the economic models. This takes economic model for granted, and estimate the parameters of such models. Most of this type of empirical work involves with several agents competing with each other, or in the case of single agent, involves with agents making inter-temporal choices. The issue here is less of which model is correct, but rather there is only one credible model so let’s take it and find out how accurate the model is.

Keane and Wolpin (1997) on estimating inter-temporal choice models of young men’s career choice (inherently dynamic, and no alternative models are available)

Estimating relationships between human capital and health capital (inherently dynamic, and no alternative models are available)