

Theory and Evidence as Drivers of Economists' Opinions Regarding the Impact of Fiscal Stimulus

Working Paper

Edward E. Leamer¹ and Sumit Shinde^{2,3}

November 16, 2020

We explore in this paper how the opinions of economists regarding the impact of a fiscal stimulus have evolved over the last century. In describing the evolution of economists' opinions, we contrast deductive methods based on words, pictures and math including calibrated models versus inductive methods using data images, tables of statistics and estimated models.

In Section 1 we contrast fiscal stimulus with fiscal policy more generally. We use the words "fiscal stimulus" to refer to policies enacted in recessions or soon thereafter, intended to reduce the amount of slack in the economy. The target of a fiscal stimulus is the unemployment rate but the fiscal GDP multiplier is the usual measure of effectiveness.

Section 2 suggests that the organic human system we call an economy has a healthy state, an unhealthy state and periods of recovery from the unhealthy state. This metaphor contrasts with the mechanical metaphor that controls much of the discourse in macroeconomics referring to the "laws of motion" of the system and to "shocks" that drive the economy into recessions. With the illness metaphor in place, we briefly describe the theoretical basis of medieval medicine before the microscope and before clinical trials, which we suggest is analogous to current macroeconomics.

Section 3 is a look at the critical role of economic theory in the formation of beliefs of economists beginning with the simple one-sector supply and demand model described in words, then in pictures and last in mathematics. Math has become the language of economists, a language that makes economists think they are scientists because Physics uses math too. We argue that is important to distinguish the mathematical properties of a model from the messages it is intended to convey to economists. The proof of the existence of general equilibrium was heralded as a final proof of the efficacy of the market system when it's only a mathematical result devoid of economic messages. A message of that model might be a way to characterize economies that have market inconsistencies and what policy might do about that. Empirical work that takes the model literally without any enhancements to capture other aspects of reality is testing the math not the message.

¹ UCLA, Anderson School of Management

² UCLA, Department of Economics

³ We thank Jesper Böjeryd for helpful comments and conversation.

Section 4 explains why membership in the two theoretical tribes (interventionists and non-interventionists) is so little affected by empirical studies: “Persuasive evidence is hard to find.” While econometric evidence in principal could be persuasive, the need to make the model capture most of the relevant forces inevitably overloads the data with too many parameters to estimate, leaving the conclusions quite fragile: change the assumptions a bit, and something very different will emerge. We argue that calibrated models should be thought to be numerical theorems not empirical evidence of policy effects, going a bit farther in criticism than Hansen and Heckman(1997).

Section 5 describes how economists’ beliefs about fiscal policy evolved mostly because of external events including the Great Depression, the Inflation of the 1970s, the Great Moderation and the Great Recession. Section 6 offers some thoughts regarding how a new kind of empirical work might help understand the persistence of high unemployment levels and the fiscal policies that might matter.

1. Fiscal Policy and Fiscal Stimulus

The Federal Reserve Act of 1913 established the Federal Reserve Board (FED) and legislatively defined monetary policy by instructing the FED to “maintain long run growth of the monetary and credit aggregates... so as to promote effectively the goals of maximum employment, stable prices, and moderate long-term interest rates.” But unlike monetary policy no federal act defining fiscal policy exists. The Act of Congress Establishing the Treasury Department of 1789 tasked the Treasury with managing government revenue and expenditure. Neither this act nor any subsequent act defines what fiscal policy consists of and what its objectives are. In the US, fiscal policy is a product of deliberative efforts of Congress and the executive branch. Monetary policy is largely separated from the political process but fiscal policy is largely controlled by politics.

In the academic literature and in public policy discussions, fiscal policy covers the full range of spending and taxation alternatives available to the governments (federal, state and local) which affect the economy. Fiscal policy during economic expansions when the market system is working smoothly has the goals of efficient provision of valuable government services, strong overall economic growth and a fair distribution of income including government benefits. Inflation is not a direct target, but can be an indirect one if inflation jeopardizes the other goals. Borrowing during expansions is appropriate when it is used to finance government investments that help to grow the future economy, such as infrastructure or winning wars or education. That kind of borrowing wisely puts some of the costs on those who will benefit: future taxpayers. Borrowing during expansions to pay for current expenses like Medicare and Social Security might fall under the heading of “a fair distribution of income” if the US Congress regards it fair to put high taxes on future taxpayers to take care of our elderly today.

This document is about “fiscal stimulus” not “fiscal policy.” We use the words “fiscal stimulus” to refer to special federal government spending and taxation enacted in recessions and their aftermaths when the rate of unemployment is elevated for an uncomfortably long period of time. A fiscal stimulus is intended to reduce that rate of unemployment. A reduction in the unemployment rate creates a larger than normal increase in employment and thus greater economic growth but growth without a reduction of the unemployment rate is not the target.

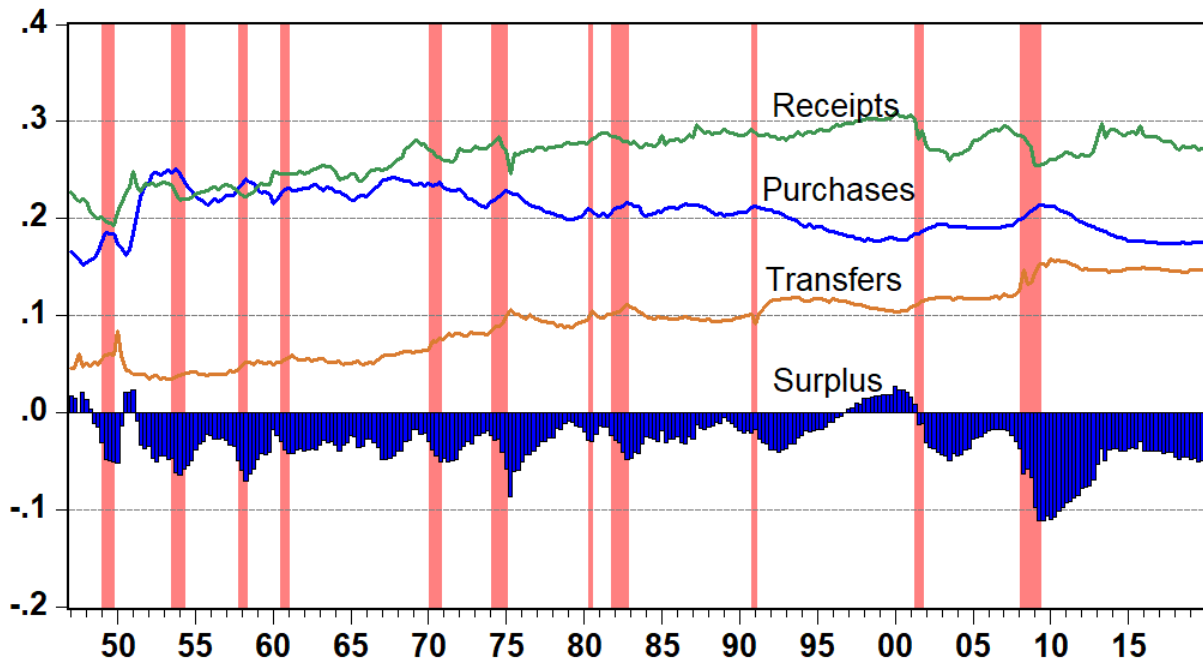
A fiscal stimulus is a combination of a cut in taxes and an increase in government spending. Government spending can include transfers to individuals including pre-committed or supplemental unemployment benefits, and also transfers to State and Local Governments which have balanced

budget requirements that force spending reductions when revenues decline in recessions. The pre-committed transfers are usually called automatic stabilizers because they increase in recessions when fiscal spending is needed to grow GDP and to reduce unemployment.

1.1 A Look at Fiscal Policy since 1947

The figure below depicts Government Receipts, Purchases (G in the GDP Accounting Identity), Transfers and the Fiscal Surplus for all levels of government as a share of GDP. Purchases as a share of GDP have been on the decline since 1970, but that has been offset by the rise in transfers. Together these two exceed Receipts from 2001 onward and create an overall fiscal deficit that has to be financed by borrowing.

US Government Receipts, Purchases and Transfers / GDP
 Recessions Shaded: Peak+1 to Trough
 Includes State and Local Governments



The shaded regions in the figure are the US recessions, and it is apparent that all four of these items are affected by the recessions: lower receipts as a share of GDP, increased purchases, increased transfers and an increased deficit. Bear in mind that between 1945 to 2019, in 75 years, US has experienced only 11 recessions which highlights the limited availability of data to study the impacts of fiscal stimuli. We have done some econometric work that identifies the increase in the *unemployment rate*, not the level, as the recession driver of these four items, with transfers being the most affected. It is important to take note of the fact that most fiscal stimulus studies deal with the fiscal multiplier applying to purchases (the fiscal purchases multiplier) and a few deal with receipts (taxation) and even fewer with transfers. This may be the result of the fact that the transfers such as unemployment benefits are mostly determined by earlier legislation and the preset rule makes it almost impossible to determine how the economy would have behaved without the “automatic stabilizers.”

The unusual rise in government purchases in the Korean and Vietnam wars is evident in this figure, as is the buildup at the end of the Reagan presidency before the end of the Cold War. These events are sometimes treated as “randomized experiments” for the study of the efficacy of fiscal policy, but none of these increases in spending occurred during recessions when a fiscal stimulus might be deployed. The sharp drop in spending when the Korean War ended did coincide with the beginning of a recession suggesting that a negative fiscal stimulus can cause a recession. The correct lesson learned from this event is probably that “overbuilding” can make the economy fragile. We are referring to the unsustainable buildup of spending on wartime manufacturing products, and when the war ended it precipitated a huge transfer of productive resources from defense to the private sector leaving unemployed resources until new work was found.

What is left out of this figure is counter-cyclical regulatory policy which can also help to control the business cycle. 1.2 The Fiscal Multiplier

Many economic models express fiscal and monetary policies with three variables G for spending, T for taxes and r for interest rates. Real life is much more complex. In monetary policy the instruments that central banks have are several- namely, the discount rate, reserve requirements, open market operations, and the interest on reserves [[Federal Reserve Bank of St. Louis Website](#)]. In contrast, under the heading of fiscal policy, though there are three classes of policy instruments (purchases, transfers and taxes), each one contains a wide range of alternatives that vary by individual, by industry, by geography, and by form. For example, the government can provide individual tax relief by giving lump-sum rebates, work pay tax credit and payroll tax holiday, reduction in individual income tax rates, or reduction in dividend and capital gains taxes. Or the government can provide business tax incentives by reducing the corporate income tax, providing incentives for investment and hiring or allowing net operating losses to be carried back [Congressional Research Service Report R45780 (2019)]. These are not likely all to have the same impacts. Furthermore, even the same instrument is not likely to have the same impact in different time periods.

The impact of a fiscal spending stimulus is generally summarized by a statistic called the fiscal multiplier. The fiscal multiplier is an answer to the question “For every one dollar increase in spending⁴ by the government what is the dollar increase in the total output?” The debate about whether government should intervene or not, essentially, boils down to this statistic. With full employment and with government spending financed with taxation, fiscal spending transfers resources from the private to the public sector, which is a good thing if the public value created exceeds the private value lost. The dollar amount measures the private sector loss, and the gain has to be the same amount or more to justify the policy. It is this kind of thinking that supports the conclusion that if the fiscal multiplier is bigger than unity then the government’s intervention is considered efficacious, otherwise a multiplier less than one implies that the fiscal stimulus should not be deployed. However, reliance on GDP growth to determine the impact of a fiscal stimulus is endangered by the fact that the value of the government contribution to real GDP is not easily determined because there is no market-based price of government services.

The Keynesian deductive story, which was the first theory to ascertain that multiplier is bigger than unity, is that the government can hire one group of unemployed workers to dig ditches and another to fill the ditches in. Nothing of value would be created but it would put income into the pockets of the

⁴ This spending can be a transfer, purchase or a tax cut.

formerly unemployed workers. When these workers spend that income, some of the spending pays for other unemployed workers to do some valuable work. Suppose that 50% of the income serves this purpose. When that 50% is passed to the next group of unemployed workers they would spend 50% and so on. So in terms of income created it's the infinite sequence $1 + .5 + .5^2 + .5^3 + \dots = 1 + c/(1-c) = 1/(1-c)$ where c is the spending fraction called the marginal propensity to consume. While the first dollar spent didn't create any valuable output, the additional spending created by this dollar was for goods or services of value. Thus each dollar of government spending created $1/(1-c)$ in total income, and $c/(1-c)$ in economic value. Incidentally, the relevant marginal propensity to consume should probably refer to purchases of the goods and services produced by the unemployed, not the total spending.

Furthermore, this focus on the multiplier misses the income distribution target of a Keynesian stimulus which is intended to create enough demand to put unemployed workers and other inputs back to work. A multiplier less than one can be justified if the spending comes with a significant decline in unemployment and a reduction in income inequality.

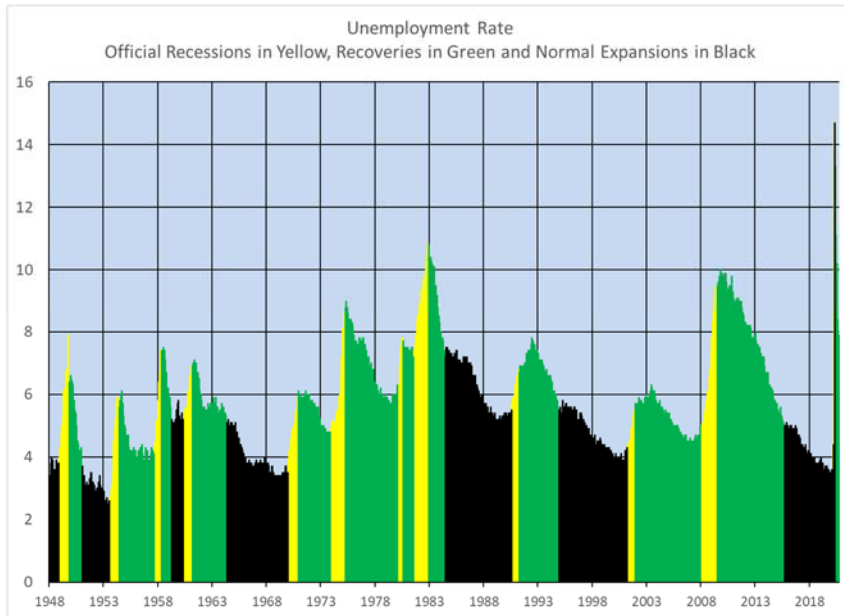
2. The Illness Metaphor

In the standard DSGE models used to study business cycles, the "laws of motion" of the economy are built into the model and recessions are assumed to be the result of exogenous shocks to total factor productivity, commonly called "productivity shocks", a mechanical metaphor. We propose the alternative organic metaphor that the economy is a biological organism that enjoys healthy periods but also diseased stages and recoveries. Monetary and Fiscal medicine are routinely used to treat this patient. This disease metaphor can affect both the deductive and the inductive approaches for studying fiscal stimulus.

The US economy has two "official" states usually called expansion and recession. The definition of a recession on the NBER website is multifaceted:

"The NBER does not define a recession in terms of two consecutive quarters of decline in real GDP. Rather, a recession is a significant decline in economic activity spread across the economy, lasting more than a few months, normally visible in real GDP, real income, employment, industrial production, and wholesale-retail sales."

Rather than a list of data series used to date a recession in an ambiguous way, we prefer to define a recession as a sustained period of elevated unwanted idleness of labor, capital and land. Idleness of labor is illustrated in the unemployment data in the figure below which has the official NBER recessions colored yellow. Here we see that sharp increases in unemployment almost perfectly replicate the NBER recession periods. Where there is a conflict, we may ask which is better?



The definition of a recession that refers to unwanted idleness is motivated by a disease metaphor for the economy: a healthy economy in which the level of idleness is “normal” and an unhealthy economy with defective markets that allow sharp increases in idleness. The disease metaphor suggests a third recovery phase during which the economy returns to “normal”. In pursuit of a definition of recovery, the figure above has months in green during which the unemployment rate exceeded the unemployment rate at the previous peak. Among other concerns, this creates five expansions that never left the recovery phase, though in four cases the unemployment rate had stabilized, suggesting a “new normal.” An alternative used in Leamer(2020) is to define the recovery period as a return to previous peak payroll employment.

Questions about remedies: The disease metaphor with three possibly-related stages suggests three fundamental questions regarding macroeconomic policies:

- What “medicine” can be administered *during expansions* that would lengthen the expansions without making the recessions that occur longer and/or more severe?
- What “medicine” can be administered *during recessions* that would shorten the recessions and make them less severe without making them more frequent?
- What “medicine” can be administered during recoveries, that would shorten the recovery without making the next disease outbreak sooner and more severe?

Questions about causes: Before seeking an answer to the three medical questions, we may need to make some scientific progress analogous to the building of the microscope and the discovery of bacteria that allow us to answer two foundational questions:

- What is going wrong with markets in recessions that allows/causes the level of unwanted idleness to zoom upward?
- Why doesn’t the unemployment rate quickly return to its pre-recession “normal” level after a recession ends? Why does the rate instead experience a pretty steady decline during the expansions?

Adopting a health metaphor would have many favorable effects. The metaphor exposes the fact that most of economic studies that estimate fiscal multiplier rely on exogenous variations in fiscal spending that arise independent of the state of the economy. As Parker (2011) pointed out that these studies answer the question that “how fiscal policy behaves on average across these 3 states”.

Second, the traditional DSGE based fiscal stimulus question focusses exclusively on policy during recessions which arise due to exogenous changes to productivity, and the models do not allow one to explore policies during expansions that can create a fragile economy and raise the chances of a recession soon. To highlight this, two examples are in order.

One example is Leamer (2008) who argues that the Fed interest rate policies affect the timing of the building of homes and not the totals. Exceptionally low interest rates in expansions take home building from the future and make the future fragile. Thus the time to increase interest rates is in the middle of expansions when US housing starts are already above the normal 1.5 million per year and growing larger. The Fed’s view is that it cannot stop bubbles from forming but it can clean up the mess that bursting bubbles create by deploying low interest rates in recessions. But when it comes to housing, during a recession when home prices are declining, the Fed is virtually powerless to stimulate home building because the negative appreciation makes homes quite unaffordable even with zero interest rates. The Fed can prevent the mess from occurring, but cannot clean it up.

Another similar point is that fiscal policy and especially automatic stabilizers can make the recessions worse and more frequent. This occurs because the price of cyclical risk is borne by the government and not by the consumers of the products that the cyclical industry is producing. A cyclical industry would ideally compensate its workers through private channels for the cyclical risk. So either a factory worker gets the unemployment benefit from General Motors or gets high rates of pay before the economy collapses. If the economy relies heavily on the public sector for fiscal stimulus, then the auto workers recognize they basically have two employers. One is General Motors when the economy is going well. The other one is the US government when things are not going well. This can mean too many new cars in expansions and deeper and longer lasting recessions. It’s like building a home on a flood plain because the government will take care of you if a flood occurs.

In addition, regulation of financial markets is an important policy that should be deployed in expansions to help make the recessions fewer and less severe, and counter-cyclical regulation might help to shorten the recessions. For example, regulation needs to control the leverage of important banks to assure their solvency and needs to limit the distribution of financial instruments that create systemic fragility like credit default swaps. It also bears saying, even if it is obvious, the medicine for a recession should depend on the character of the expansion that preceded it and the nature of the sectors that are demand-constrained. It’s a complex ever-changing intertemporal problem not adequately captured by traditional macroeconomics theory.

2.1 A Brief History of Medical Practice in the 19th Century

With the illness metaphor now in place, it seems useful to explore how human illness was treated before the discovery of the role of bacteria.

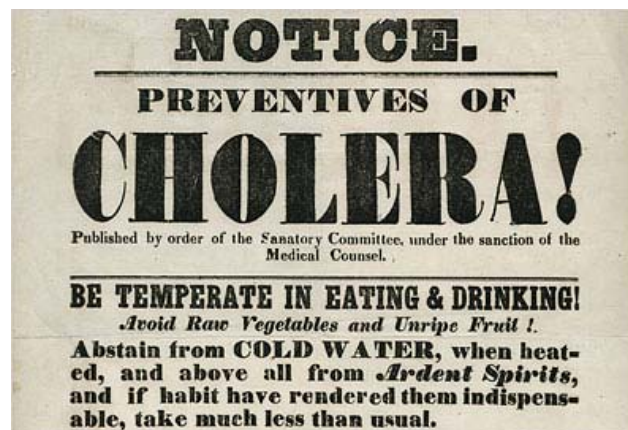
When plagues and pandemics swept through Europe in the middle ages, two medical theories emerged: the miasma theory and the contagion theory. These theories recommended different responses: either clear the air, or socially distance.

Faced with an onslaught of terrible diseases in the Middle Ages, the people demanded cures. With little or no understanding of the causes of disease, remedies still had to be offered and they were, lots of them.

Medieval medicine was practiced by two distinct groups:

- (1) The theoreticians who based their remedies on reasonings of church officials and the ancient Greek medical philosophers, Hippocrates and Galen.⁵
- (2) The empirics who based their remedies on their own experiences. The first definition of an empiric is: one who relies on practical

Hand bill from NYC Board of Health, 1832



https://en.wikipedia.org/wiki/Cholera#/media/File:Cholera_395.1.jpg

⁵ “The ancient Egyptians developed the theory of humorism, Greek scholars and physicians reviewed it, and then Roman, medieval Islamic, and European doctors adopted it. Each humor was linked to a season, an organ, a temper, and an element.

Humor	Organ	Temper	Season	Element
Black bile	Spleen	Melancholy	Cold and dry	Earth
Yellow bile	Lungs	Phlegmatic	Cold and wet	Water
Phlegm	The head	Sanguine	Warm and wet	Air
Blood	Gallbladder	Choleric	Warm and dry	Fire

The [theory](#) held that four different bodily fluids — humors — influenced human health. They had to be in perfect balance, or a person would become sick, either physically or in terms of personality.”

<https://www.medicalnewstoday.com/articles/323533.php#middle-ages>

experience. The second definition reveals the political power of the theoreticians: a charlatan, a quack.

Theorists divided themselves into two tribes: the miasmatis⁶ who believed disease was spread by putrid air (miasma) and the contagionists who thought disease was directly passed from person to person. During the 1854 London cholera epidemic, the miasmatis Thomas Calley suggested clearing the air by firing bags of gunpowder from cannons. Contagionists counseled people to move to the countryside where human contact was less. Meanwhile, John Snow discovered that exceptional rates of cholera in London were suffered by those who took their water from a Broad Street public water pump. This finding did not fit with either of the prevalent stories, and absent a compelling story it was ignored by the public health officials, who may have been inundated with many different “compelling” pieces of evidence but no stories why that evidence was anything more than just another coincidence discovered by the 19th Century version of Artificial Intelligence: incessant search for patterns without stories.

This uncomfortable equilibrium gave way to modern medicine because of the invention of the microscope by Zacharias Janssen around 1590 and the discovery of bacteria by van Leeuwenhoek in 1676 and the work of Louis Pasteur and Robert Koch advocating the Germ Theory and backing it with scientific evidence in the 1850s and thereafter.

Like the miasmatis and the contagionists, economists divide themselves into two theoretical groups: the conservative free-market advocates and the left-leaning advocates of government interventions including fiscal policy and minimum wages, for example. We patiently await the arrival in the Economics community of the equivalents of Janssen, van Leuwenhoek, Pasteur and Koch.

3. Economic Theory Controls the Conversations

3.1 The Simple Supply and Demand Model Illustrates the Power and the Limits of Mathematics as a Deductive Tool

Before we get to the evolution of opinions about fiscal policy, we make some remarks about economists’ opinions broadly. The opinions held by economists are created and communicated principally with deductive reasoning based on simple first principles such as “individuals respond to financial incentives.” From assumptions like these together with undisclosed institutional assumptions, powerful conclusions about public policy are drawn. This began famously with the Adam Smith argument that the Invisible Hand of the market turns greed into a force for good. Best, then, to keep the government out of the marketplace. No fiscal stimulus, please.

The explicit behavioral assumptions and the implicit institutional assumptions that economists routinely deploy describe an extremely orderly world. For example, the supply and demand market model that is the foundation of most economists’ beliefs captures greed and desire with precise hypothetical curves

⁶ Sandra Hempel, *The Medical Detective; John Snow, Cholera and the Mystery of The Broad*
<https://books.google.com/books?id=ssgTAAwAAQBAJ&pg=PT29&lpg=PT29&dq=%22cannons%22+to+clear+the+air+during+cholera+epidemic&source=bl&ots=fqjV-Tlyzi&sig=ACfU3U0KVjtnhbuPpQWgF4fSALe0TtQ9MA&hl=en&sa=X&ved=2ahUKEwi9jKL2sprmAhXhOn0KHUmVBCgQ6AEwAAnoECAsQAQ#v=onepage&q=%22cannons%22%20to%20clear%20the%20air%20during%20cholera%20epidemic&f=false>

that are assumed to represent the willingness of buyers and sellers to exchange one product or service for promissory notes (money) that can be spent in other markets. The market price is the value in promissory notes that exactly balances the total willingness to buy with the total willingness to sell. The supply and demand curves are assumed to be precise representations of human feelings which are implicitly assumed to last long enough to allow a “Walrasian” auction to take place so that all exchanges can and do occur at the market price. Moreover, these personal feelings that determine the supply and demand curves are “structural” meaning that they are not influenced at all by the Walrasian auction. The fact that you are willing to pay a high price does not affect my opinion about the value of the item.

While we might worry about the assumption that the willingness to buy and sell reveals values accurately, it’s the usually unspoken institutional assumption of a Walrasian auction that really seems completely fanciful. For example, economists routinely refer to the labor “market,” and leap from those words to the unannounced assumption of a Walrasian auction, when in reality the labor “market” is a series of bilateral negotiations and renegotiations, each of which is influenced by the results of other negotiations but few if any occur at the unknown “market” price.

One Keynesian contribution to macroeconomic theory used to justify fiscal policy is the assumption of downwardly inflexible nominal wages in the labor market supply and demand model, but if one thought of the labor contracts as a sequence of bilateral bargains, one would expect a ‘price discovery’ process during which the realized contracts would drift slowly from one market equilibrium price to another. Each renegotiation at the end of an existing labor contract would be anchored by the terms of the previous one and in a temporary recession one would expect at most a slow rate of descent in average wages of continuously employed workers. Focus might be put on the contracts formed between unemployed workers with their previous employer versus a new employer, one anchored by a past contract between both parties and the other with an employee-sided anchor of the characteristics of the employee’s previous job and an employer-sided anchor based on other recently formed contracts. While the assumption of a market equilibrium with a wage floor might be a useful approximation to this price discovery process, this fixed wage theory says nothing about how the fixed wage is determined and how it changes over time.

Sales of homes, like labor market contracts, occur in a sequence of bilateral bargains. As Leamer (2008) argued, housing plays an important role in the US business cycle. One reason for this is that sales of existing homes in recessions are greatly influenced by the ability and willingness of current owners to hold on to their homes until the housing prices recover to the point where sellers can get a “fair” price again. The few sales that occur are at the sellers’ prices not the market prices. This creates what Leamer(2008) refers to as a volume cycle not a price cycle, with sharp declines in sales of homes but only modest and sluggish price declines. However, the 2008/09 recession had sharp home price declines because banks became owners of a large number of homes with defaulted subprime mortgages, and because of accounting rules banks were highly motivated sellers and they dumped those homes onto the market with little concern about the prices that could be negotiated. This is another case where a price-discovery theory suggests studies of actual data to lend it evidence-based credibility.

Macroeconomics, in our opinion, should explicitly allow for price discovery processes and not rely on the Walrasian assumption. One could imagine that an economy has a healthy phase during which price discovery is orderly and rapid and an unhealthy phase (recession) during which price discovery is sluggish, hanging on to pre-recession prices. For example, when the Department of Defense cut back

spending in response to the collapse of the Soviet Union, it took a long time for aerospace engineers who built laser-guided weapons to come to grips with the reality that their next best alternative was driving taxis. Also it seems likely that the price discovery operates differently in an inflationary environment than a deflationary environment.

3.2 Mathematics Has Become the Language of Economists

Following a course in microeconomics, many exams in economics have asked students to determine what happens to the market price and market quantity if demand and supply are linear and the demand increases while supply is held fixed. That is a mathematics problem involving the solution of a two-equation system. An economics question might ask for an essay explaining: What might happen to the price of rice if the Mississippi river floods? The mathematics question has a precise correct answer that can be graded by a computer, but the economics question has no precise correct answer and is difficult to grade. That makes math a great language in an economics course – it's easy to grade.

Mathematics is also a great language for uncovering logical errors but it comes with the seductive appearance of precise conclusions even when the assumptions are only approximations chosen for rhetorical purposes. Samuelson, more than any other economist, was responsible for importing math into economics, most notably with his book *Foundations of Economic Analysis* published in 1947, which turned Economics into a set of straightforward constrained optimization problems: maximizing utility given resource constraints. When Samuelson tried out this new language with a submission of his article with Wolfgang Stolper to the *American Economic Review*, he was rebuffed by the editor Homan(1941): "It is a very narrow study in formal theory, which adds practically nothing to the literature of the subject with which it is nominally concerned." The paper was published later in the same year by the *Review of Economic Studies* which apparently was more receptive to the mathematical language.

Was this paper a contribution to economics, or not? Here is a surprising conclusion of the paper:

It is the intention of the present paper to show that definitive statements are possible concerning the effects of international trade upon the relative remunerations of productive agencies, and more important, upon their absolute real incomes... In the beginning we expected to do no more than delineate factors which would indicate likelihood in one direction or another, only in the course of the investigation did we discover that **unambiguous inferences** are possible. (Stolper and Samuelson 1941, p.334) Leamer (27)

Many, notably Bertil Ohlin, the namesake of the Heckscher-Ohlin general equilibrium model, had expressed written ideas in support of the view that everyone can gain from international trade. The "unambiguous inferences" uncovered by Stolper and Samuelson are that one "productive agent" has an increase in real incomes and the other has a reduction. But the use of the word pair "unambiguous inferences" is astonishing rhetoric that portends the shift in the economics discipline to come. An inference is usually defined as a conclusion that is drawn from evidence and reasoning. Stolper and Samuelson refer to no evidence. And the reasoning that is offered is mathematically correct but highly misleading. The Stolper-Samuelson Theorem refers to a world with two traded products with prices determined in the international marketplaces, with two factors of production and with constant returns to scale in both productive activities. That does not seem like a good approximation of the real world, and there are lots of ways to alter the assumptions to get a different conclusion. For example, if countries have sufficiently different factor supplies, they can specialize in non-overlapping sets of

products, and everyone can gain from trade. What we are suggesting is that the economics of the Stolper-Samuelson theorem comes not from its mathematical validity but instead from the fact that it leaves us to considering the myriad of other assumptions that would lead to different conclusions. In addition, this theorem suggests a mechanism by which arbitrage in the global goods market that sets product prices translates into indirect arbitrage in the local factor markets even if these factors cannot move between countries and even if there is no movement of factors between sectors which have fixed input technologies. Perfect arbitrage is accomplished by mobility, not by movement.

This illustrates an important idea: a model has mathematical properties and it has messages. These are not necessarily the same. If you do not understand the difference, you are a mathematician, not an economist.

Here is a blurb on a book by Leamer:

Claiming “a lifetime relationship with Heckscher–Ohlin,” Leamer argues that Bertil Ohlin's original idea offered something useful though vague and not necessarily valid; the economists who later translated his ideas into mathematical theorems offered something precise and valid but not necessarily useful. He argues further that the best economists keep formal and informal thinking in balance. An Ohlinesque mostly prose style can let in faulty thinking and fuzzy communication; a mostly math style allows misplaced emphasis and opaque communication. Leamer writes that today's model- and math-driven economics needs more prose and less math.⁷

Theoretical models have math and messages in them. The math part is the arithmetic acrobatics that theoreticians do to demonstrate that the model internally coherent. These arithmetic acrobatics are phenomenally impressive and have ever been growing in complexity. As our computation power grows so does our ability to explore multiple complicated channels of influence simultaneously, understand their interaction, mute everything else and, what is more, fit all these facets in a model that is internally consistent. But in the context of study of fiscal stimulus, the channels and their interactions that influence the size of the multiplier comprise the message of the model which is often missed. This message should motivate empirical enquiries which instead are mostly focussed on the size of the multiplier.

3.3 Existence of General Equilibrium

Now that we have come this far, we need to go back to the first couple of paragraphs in this section that discuss the supply and demand model. What was not mentioned in those paragraphs is that together with the one market under study, the model must have markets for all other products which determine the value of the promissory notes that are exchanged in the market that is focused on. There must be a grand Walrasian auction for all the markets at the same time, and all the participants must know the market prices of all the auctions before they make any actual exchanges. This is an implicit assumption of most macroeconomic models.

Here is where economics has clearly turned into mathematics. Instead of expressing doubt about the institutional and informational assumptions of this general equilibrium model, mathematical economists (the high priests) asked if the Walrasian set of equations that describe the full equilibrium has only one

⁷ <https://mitpress.mit.edu/books/craft-economics>

solution, or more than one, or none at all. The fascinating book by Duppe and Weintraub (2014), *Finding Equilibrium*, describes the search for a mathematical solution to this complicated question by Kenneth Arrow, Gerard Debreu and Lionel McKenzie.

But what are the economics questions? If there are multiple equilibria, why is that, and what policies might affect the outcome? Exactly what might it mean in the real world not to have an equilibrium? Here is a possibility. If the production possibilities curve takes the form $xy < 1$ and the goal is to maximize $x+y$, there is no solution since infinity is a destination not a number. The real world in this case would want to choose a very high value of either x or y , and a low value of the other, moving more and more toward one of the corners over time. Thus although there is no mathematical solution, there is a pretty good story to tell about how this kind of economy would evolve, moving incessantly toward one of the corners, but never getting there.

The economics questions were completely abandoned in the mathematical search for an equilibrium, except in one sense. Some economists thought that the classical view of Adam Smith regarding the efficiency of a market system was at risk if the set of supply and demand curves that describe the markets did not have an equilibrium. There was therefore a huge celebration when Arrow, Debreu and McKenzie proved there was an equilibrium. Our view is that for the design of fiscal policy this mathematical result is quite irrelevant.

The assumptions of a mathematical model are approximations. Without saying so very often, economists think that their theorems are certainly true given the assumptions but still insightful in a domain of assumptions larger than the ones deployed to prove a theorem. To put it another way, we think and we hope that our theorems are robust in the sense that minor tinkering with the assumptions do not lead to major changes in the conclusions. If that hope is not fulfilled, for example in the case of the Stolper-Samuelson theorem, the meaning of the theorem derives from how it is violated.

4. Persuasive Evidence is Hard to Find

With the exception of the Phillips curve there has been little impact of empirical studies on the deep beliefs of macroeconomists and for that matter the impact of Phillips didn't last long as the curve mostly unraveled with new data. Like medieval doctors, absent persuasive evidence practitioners rely on theory.

4.1 General Issues with Inductive Evidence

Econometric theory offers two important warnings about inductive inferences from economic data. Both warnings make the same point: It takes untestable assumptions to draw causal inferences from non-experimental data. This makes many inferences drawn from nonexperimental data uncomfortably fragile, meaning they depend very much on doubtful, untestable assumptions. That's the basic reason why there is so little impact of empirical studies on the thinking of macroeconomists about fiscal policy, something anticipated by John Stuart Mill(1836), discussed below.

The two warnings are:

- (1) Omitted variable bias. A simple correlation between the target variable y and the treatment x does not reveal a causal effect of x on y because that correlation may be a consequence of some other effects w that influence both y and x . A regression of the target y on the treatment x provides biased estimates of a causal effect of x on y *unless* there are "enough well-chosen"

control variables w included in the equation. There is no reliable “specification test” to determine if the controls are adequate. Moreover, economic theory concentrates on the effect of x on y and rarely provides any advice about what controls w are needed to study actual data. An analogy is the law of gravity that describes what happens in a vacuum, but this theory provides no advice about how to study the force of gravity in a hurricane. The greatest problem for the analysis of economic data is how to translate a simplistic mathematical theory into an empirical model suited to the complexities of the real world and not to overwhelm the limited data with so many unknown parameters that hardly any can be reliably estimated. This specification problem is all the more complex if the causal effect of x on y depends on the controls w , meaning that the model has interaction variables as well as separate linear effects. Since the extent of the parameterization has to be controlled, the final parameterization either can be chosen behind closed doors, or the full parameterization can be revealed to the audience together with something like a Bayesian prior distribution that limits the effective number of parameters, as recommended by Leamer(1978). A “prior” is almost always involved, often uncovered during exploratory study of the data, but hardly ever revealed to the audience.

- (2) Two-way causality bias: Controls are not enough if the target y and the treatment variable x are jointly determined like the market price and the market quantity. Then an “instrumental variable” is needed. A proper instrumental variable z must have no direct effect on the target variable y and it must influence the treatment x . Then, a correlation between the instrument z and the target y , establishes the causal impact of w on y operating through x . The size of the causal impact is a ratio of coefficients from two regressions, y on z and x on z . A simple way to explain the logic of the ratio is to note that the causal derivative dy/dx can be written as the ratio of two causal derivatives $(dy/dz)/(dx/dz) = dy/dx$. There is no way to test the negative assumption that z has no direct effect on y . The theory of estimation of the ratio of two regression coefficients remains mostly “asymptotic” while valid inferences from real finite samples are distinctly different. In addition, the two regressions needed for instrumental variable estimation both suffer from potential omitted variable bias. Except in the case of obvious “natural experiments” like discontinuity designs, instrumental variable estimation should be treated with a high degree of skepticism, and is.

Incidentally, because of the heterogeneity of impact, we suggest replacing the word “instrument” with “surrogate” where surrogate refers to a stand-in for the actual intervention being considered. For example, the impact of education on income surely depends on how that extra education was acquired. There is likely a different effect if parents threaten to disown their children if they do not go to college than if parents agree to pay all their tuition, though both may produce the same increase in years of education. If you are thinking of paying the tuition for your child, is closeness to a university a good surrogate?

A startling fact is that the essence of the omitted variable problem was discussed by John Stuart Mill in 1836 long before it found its way into econometric textbooks.:

“Call the effect B, and let the question be whether the cause A in any way contributes to it. We try an experiment in which all the surrounding circumstances are altered, except A alone: if the effect B is nevertheless produced, A is the cause of it. Or, instead of leaving A, and changing the other circumstances, we leave all the other circumstances and change A: if the effect B in that

case does not take place, then again A is a necessary condition of its existence. Either of these experiments, if accurately performed, is an *experimentum crucis*; it converts the presumption we had before of the existence of a connection between A and B into proof, by negating every other hypothesis which would account for the appearances.” Mill(1836)

While the regression tool does not suppose that one can in the real world hold all the other circumstances fixed and allow A to vary per Mill, a regression does provide an estimate of the effect of B on A holding fixed all the other circumstances captured by the control variables. It is all the more surprising that Mill anticipates the impossibility of controlling for all the relevant variables, and endorses the method we are here calling “deductive” reasoning.

“Since, therefore, it is vain to hope that truth can be arrived at, either in Political Economy or in any other department of the social science, while we look at the facts in the concrete, clothed in all the complexity with which nature has surrounded them, and endeavour to elicit a general law by a process of induction from a comparison of details; there remains no other method than the *à priori* one, or that of “abstract speculation.”

‘By the method *à priori* we mean (what has commonly been meant) reasoning from an assumed hypothesis; which is not a practice confined to mathematics, but is of the essence of all science which admits of general reasoning at all.” Mill(1836)

With these words Mill anticipates almost two centuries of subsequent study that has left our opinions mostly theoretically determined. The only thing he missed was the need for instrumental variables.

It is possible to be more optimistic than Mill. The fundamental goal in the absence of a randomized controlled experiment is to find something that behaves like a randomized experiment. That is exactly what happens when “enough” control variables w are added to a regression. The regression coefficient of the variable x in the y given x and w model can be found by first regressing both y and x on the controls w , and then regressing the residuals of the y equation on the residuals of the x equation. Some analysts make the argument that after controlling for the variables w what is left in x is like a randomized treatment leading to an estimate that is essentially the implication of a randomized experiment. That is just another way of saying that it takes a well-chosen group of control variables to allow a valid causal inference based only on the part of the variability of treatment x that is not explained by the controls w . Mill(1836) suggests that this can never be achieved.

A causal inference from non-experimental data requires plenty of stated and unstated assumptions. Theory also requires assumptions but the role of the theorist is to concentrate our attention on one aspect of reality, not to build a model that is a complete description of the real world. A theory, like the Stolper-Samuelson theorem, can derive rhetorical power in part from the ways it is fragile. But a fragile econometric conclusion is useless. The exploratory approach needed to select the control variables is where the fragility is usually experienced, though this is often suppressed and only a favorite outcome is revealed. This creates a serious ambiguity problem: if you want an estimate of the fiscal multiplier in excess of one, there’s a model that can do that. If you want it less than one, there’s a model for that too. Why should I believe your results?

Here is our advice: There is no causal non-experimental evidence without theory! Context-free machine-learning algorithms that pick the highest correlations and the highest partial correlations

among thousands of variables should not be relied on to draw causal conclusions. While any one individual may not mimic machine learning algorithms, the profession as a whole surely does. This kind of data analysis suffers from two basic problems. One problem is that the final result can greatly underestimate the uncertainty that properly attaches to the finding. This has produced a large literature on p-hacking (e.g. Head et.al.(2015)) which should make us very wary of published values of t-statistics. Secondly, unless there is a story/theory that supports the sign and magnitude of the correlation or partial correlation, the finding of a correlation that passes an appropriate statistical significance threshold should be regarded as useful for prediction but not for the design of causal interventions. Unless there is a story that suggests the control variables, choosing any controls to achieve a desired outcome is not wise. It's Pattern and Stories (Leamer(2009)) that together make a persuasive causal argument.

4.2 Inductive Study of Fiscal Multipliers

The general regression specification that inductive methods have relied on is the following;

$$Y_{t+1} - Y_t = \alpha + \beta(G_t - G_{t-1}) + \gamma X_t + \varepsilon_t$$

where β is the fiscal multiplier, Y_t is the aggregate output, G_t is the fiscal spending, X_t is the set of covariates and ε_t is the unobserved factors in year t. One problem with this approach is that a "full" set of variables X_t can introduce enough additional parameters that the data are overwhelmed and cannot answer any question. There are plenty of other drivers that might affect GDP. What about the size of the polar ice cap? Why not?

In addition to the omitted variable problem, there is also a two-way causality problem. If the government anticipates that Y_{t+1} will decline, it will most likely counteract that by changing fiscal policy i.e. by increasing G_t . If the economy gets better after that, how do we know what was the effect of the fiscal policy? The economy has its own responses to make it healthy again. A basic problem with a simple-minded look at these correlations is that the sickest people take the most medicine, apparently suggesting that the medicine caused the sickness. All the recessions since WWII have been followed by periods of recovery with declining unemployment rates and exceptional GDP growth. Government policies enacted during the recession can claim credit for the recovery, just as medieval doctors did when patients recovered after blood-letting. Valid scientific confirmation of the efficacy of medicine requires a randomized trial or a compelling natural experiment.

A problem that afflicts conclusions drawn from either randomized trials or from natural experiments is that what is learned in the "lab" may not apply in other settings. For example, the intertemporal apparent randomness of wars which require a fiscal response is often treated as a natural experiment (e.g. Barro and Redlick (2010)), but the impact of a wartime ramp-up in spending is probably very different than a peacetime infrastructure bill or a tax cut. Another source of randomness in fiscal policy is politics (Romer and Romer(2010)). Other than wars and politics and pandemics, there are no appropriate "surrogates" for randomization at the aggregate level and therefore no evidence to support causal conclusions about fiscal policy, and, in addition, wars and pandemics create circumstances that allow fiscal effects to differ from what they might be in normal recessions.

The fact that fiscal policy is inherently political has helped researchers identify convincing natural experiments. LBJ waged a "war on poverty" and Bush Jr. waged a "war on terrorism". The intertemporal apparent randomness of wars which require a fiscal response is often treated as a natural experiment

for government purchases (e.g. Barro and Redlick (2010)). Similarly, not all tax cuts are in response to economic downturns. Some tax cuts were due to reasons related to political ideology such as reduction in size of government or fair income-distribution which can be treated as a natural experiment (e.g. Romer and Romer (2011)).

A structural VAR approach used by Blanchard and Perotti(2002) and others does not avoid the two basic problems of omitted variable bias and two-way causality. It addresses them directly. A structural VAR approach imposes the assumption of one-way causality from fiscal policy to GDP and it requires enough control variables so that what is left in the variability of fiscal policy is like a “random shock.”

“First, in contrast to monetary policy, fiscal variables move for many reasons, of which output stabilization is rarely predominant; in other words, there are exogenous (with respect to output) fiscal shocks. Second, again in contrast to monetary policy, decision and implementation lags in fiscal policy imply that, at high enough frequency-say, within a quarter-there is little or no discretionary response of fiscal policy to unexpected contemporaneous movements in activity.”
Blanchard and Perotti, Page 1130.

4.3 Deductive Calibration

The deductive method has what appears to be an empirical phase called calibration in which the parameters of a theoretical model are chosen so that either (1) the model’s predictions have moments that match the statistical moments of an actual data set, or (2) the parameters are set equal to their values estimated in other studies, often using microeconomic data. We consider this to be a numerical theorem, not empirical evidence. This is taking the theory to be literally true, and it involves no attempt to add the controls needed to allow the theory to be used sensibly to study complicated real world data.

Calibration can be useful if it produces unexpected outcomes, which can force a rethinking of the model’s assumptions. That is true of theory generally: the power of theory is uncovering something that is implied by the assumptions that is quite unexpected. The Stolper-Samuelson theorem is a great example. Ohlin thought that international trade increased incomes of all factors of production, but Stolper and Samuelson offered a simple theory with the surprising conclusion that there are winners and losers. This opened theory and evidence in search of circumstances in which there are only winners.

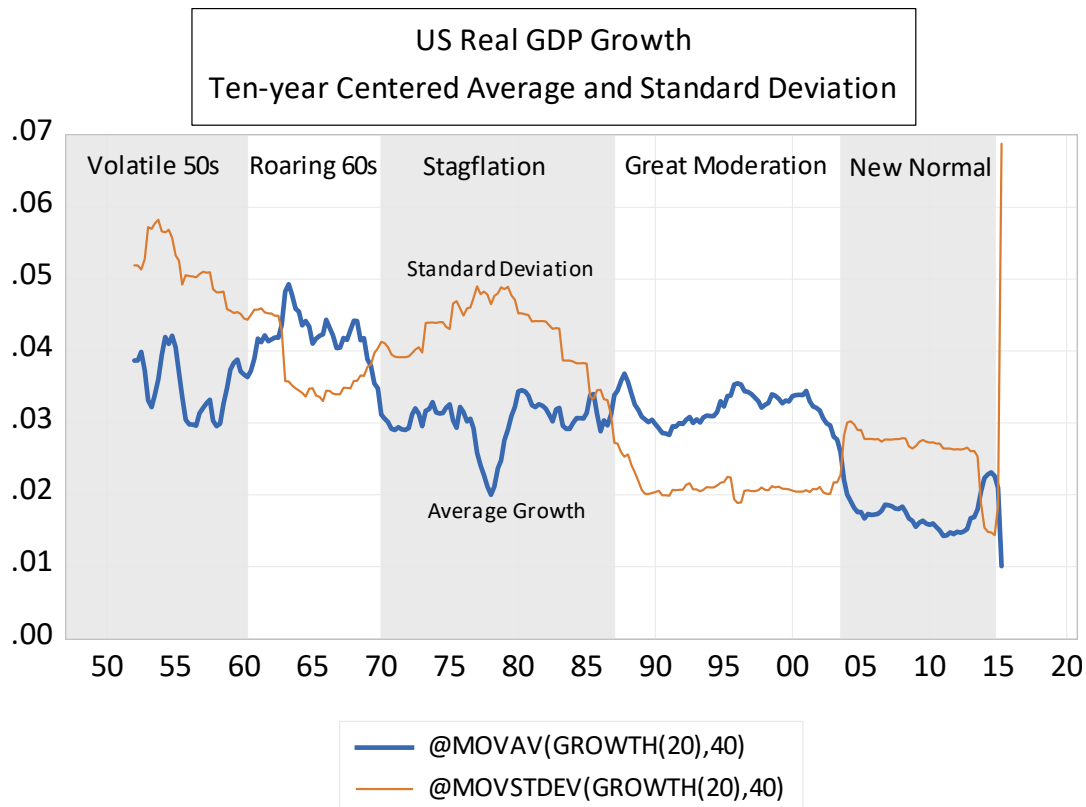
The pitfalls of calibration can be illustrated using a simple Keynesian model that imposes the assumption that the multiplier is fully determined by the marginal propensity to consume. The model can then be “calibrated” with an econometric estimate of the MPC. The resultant multiplier is not based on any direct evidence of the effect of government spending. This is taking a mathematical property of the model much too seriously. A wise economist understands that models are designed to focus on one aspect of reality and when the empirical evidence is organized it is essential to add to the model other important aspects of reality. A wise economist also understands that drawing causal conclusions from non-experimental data is a very difficult task and ultimately relies on assumptions about the causal path.

Regardless of the approach, there are two critical questions that an analyst should answer: (1) “What feature of the data makes you think your estimate of the multiplier is a good one?” (2) How reliable is your estimate? If there were an experiment, the answer to the first question would be the contrast between the treated group and the untreated group, and the answer to the second would be a standard error that depends on the difference in the means, the population standard errors and the number of

trials. The first question can be answered with a data perturbation exercise that demonstrates how answers change as the data change. It is not common for either econometric estimation or calibration to conduct a data perturbation exercise. The second question is answered with standard errors attached to estimates, which is an automatic part of an econometric exercise that has no counterpart in calibration. Calibration lacks a means of doing this. Both approaches also should do theory perturbation exercises to determine how changes in the theory change the conclusions.

5. The Evolution of Economists' Thoughts about Fiscal Policy

Views of economists about the effectiveness of fiscal stimuli have been influenced by the different circumstances depicted in the graph below, which illustrates the ten-year centered moving average growth of US real GDP and the standard deviation of growth over the same period which basically picks up the frequency and severity of recessions. The average growth in the 1950s was in the 3-4% range but the standard deviation was quite high because of the high frequency of recessions. The Roaring 1960s brought an exceptional period of growth in excess of 4% and much lower volatility. Average growth shifted down to 3% beginning in 1970 and maintained that level until the Great Recession. The standard deviation jumped up in the 1970s in a period often called Stagflation, and remained high until the great moderation that began in the later part of the 1980s. Then the New Normal kicked in after the Great Recession, with low average growth and a jump up in the standard deviation. This graph focusses the question: did fiscal policy reduce the volatility or increase the average rate of growth?



5.1 To Intervene or not to Intervene: That is the question

When economic recessions swept through the industrializing world in the second half of the 19th century, the efficient markets theory remained dominant, and its advice to governments was “Do nothing, and this disease will quickly abate. If you do something, that will make the situation worse.” Most people did not hold the government responsible for economic conditions (Appelbaum (2019)). Though there were other proto Keynesian ideas favoring government intervention supported by prominent economist such as A. C. Pigou, these ideas were vehemently opposed.

The underlying principle is that the Government should add to the effective demand for labour at the time when the effective demand of private traders falls off. The signatories of the Report took the economic soundness of the principle for granted, and did not argue it. But they found a supporter in Professor Pigou, and if their principles have gained very general acceptance, that is largely due to his high authority...The original contention that the public works themselves give additional employment is radically fallacious.

- R. G. Hawtrey, 1925

As Keynes put it “The difference is between those who believe in a self-adjusting economy and those who don’t. Those who do will always argue for non-intervention in order to allow freedom to economic factors to bring about their own self-adjustment.” He adopted the idea of a multiplier, which was first introduced by R. F. Kahn in 1931, to argue in favor of government intervention. The new theoretical foundations and long-lasting market failures during the Great Depression of the 1930s crushed the prevalent laissez-faire belief in the efficacy of the free market system, and fiscal stimulus advocates led by Keynes became more and more influential over time.

The power of these Keynesian ideas may derive mostly from the fact that they offered some medicine for the severe depression disease, while the prevalent efficient markets theory offered none, which would be ok if the disease was mild and did not last long. Best to keep in mind that when harmful diseases are inflicting humanity and no reliable treatment is available, quack cures will surely fill the need. We might as well give them a try since there is no alternative treatment. Try blood-letting or fiscal policy. Actually, in some settings blood-letting had favorable impacts. That seems likely for fiscal policy too.

There is one strictly empirical finding that had a big impact on macroeconomics thinking. It’s the 1958 Phillips curve that revealed a previously unexplained negative relationship between inflation and the unemployment rate, suggesting cost-push inflation or worker market power or who knows what. In the wake of publication of the Phillips curve in 1958, Samuelson and Solow argued that a “menu of choice between different degrees of unemployment and price stability” existed and suggested that the government could “direct the regional distribution of its expenditures more deliberately in terms of the existence of local unemployment and excess capacity” to move along the Phillips curve (Samuelson and Solow (1960)). After the apparent success of the 1964 Kennedy Tax Cuts in stimulating growth of the US economy, even the free-market advocate Milton Friedman in 1965 announced that “We are all Keynesians now.” Based on their new-found belief in the power of fiscal policy, economists at that time celebrated the end of the business cycle, prematurely, since recessions continued to occur. To quote then President Lyndon Johnson “Full employment with stable prices; Rapid economic growth; ... These

perennial challenges to economic policy are not fully mastered; but we are well on our way to their solution. (Economic Report of the President (1965))”.

As the 1973 oil price shock shook the world, the United States found itself in such a crisis that the only word that could describe it was invented just a decade earlier across the pond: Stagflation. Though unemployment rose periodically, inflation for over a decade was well over 5% reaching its peak in 1980 at 13.5%. Jimmy Carter decided to “concentrate on reducing the budget deficit by holding down Federal spending and forgoing tax reductions” (Economic Report of the President (1980)). As the economy seems to have grown unresponsive to the old medication of tax cuts and federal spending something new had to be tried. Paul Volcker was brought in to “slay the inflationary dragon” and it worked! The decline in the volatility of GDP in the 1980s was called “The Great Moderation”, credit for which was assigned to monetary policy, and celebrations commenced again, not of the end of the recessions but instead of the promise of more moderate ones that occur with wisely chosen monetary policy. In other words, the medicinal advice shifted from fiscal to monetary policy. This shift in policy was advice also mirrored in economists’ beliefs and modelling choices. Macroeconomists arrived at a consensus that models should be micro-founded and forward looking, and Dynamic Stochastic General Equilibrium (DSGE) models were adopted. Ideas such as Lucas’ critique and Ricardian Equivalence reflected the mood of the times to come and asserted that interventionist policies were close to futile.

From 1965 to the precipice of the Great Recession, economists’ belief in the efficiency of the free market became very much reinvigorated, so much so that Fed Chair Alan Greenspan ignored the warnings from Ted Gramlich in 2000 that subprime lending was creating large numbers of highly dangerous home loans and Treasury Secretary Larry Summers worked against the regulation of markets for derivatives, financial products that Warren Buffet had called Weapons of Mass Destruction. “Let the financial markets do their magic, without government encumbrance” was the meme. Being unaware of the precarious financial risk that was growing, president Bush, in February 2007, wrote to Congress that “Our economy is on the move and we can keep it that way by continuing to pursue sound economic policy based on free-market principles. Sound economic policy begins with low taxes.”

The Great Recession of 2008/09 might have delivered a shock to economists’ belief systems comparable to the Great Depression. But the actual persistence of the beliefs of economists is revealed by two opposing letters written in January 2009 in the midst of the recession. The first below, on the left, called for a fiscal stimulus and the second below, on the right, argued against it. This is the economics equivalent of the miasma and the contagion theories of disease: sensible ideas with little evidential basis.

Letter to Congress:

Economists Across the Spectrum Endorse Stimulus Package

January 27, 2009, 9:00 am

We the undersigned encourage Congress to pass quickly the American Recovery and Reinvestment Act of 2009 and stem the tide of rising joblessness. The United States is in a recession that threatens to be deep and protracted. Each month, employers are shedding hundreds of thousands of jobs. To stop the hemorrhaging of jobs and pull the economy back from the edge, policymakers must act quickly and decisively. **A critical needed action is significant fiscal stimulus specifically designed to boost employment and economic growth.**

“There is no disagreement that we need action by our government, a recovery plan that will help to jumpstart the economy”

Signed by over 200 economists

“There is no disagreement that we need action by our government, a recovery plan that will help to jumpstart the economy”

President-Elect Barack Obama, January 9, 2009

With all due respect Mr. President, that is not true.

Notwithstanding reports that all economists are now Keynesians and that we all support a big increase in the burden of government, we the undersigned do not believe that more government spending is a way to improve economic performance. More government spending by Hoover and Roosevelt did not pull the United States economy out of the Great Depression in the 1930s. More government spending did not solve Japan’s “lost decade” in the 1990s. As such, it is a triumph of hope over experience to believe that more government spending will help the U.S. today. **To improve the economy, policymakers should focus on reforms that remove impediments to work, saving, investment and production.** Lower tax rates and a reduction in the burden of government are the best ways of using fiscal policy to boost growth.

Signed by over 200 economists

Cato Institute

In last 100 years, a phenomenal amount of innovation has occurred in the realm of economic models. Nonetheless there seem to be two separated tribes: *no-intervention* and *pro-intervention*. Each tribe is taking turns to sway the policy makers in their favour. As Keynes puts it “But just because the differences go so deep there is no chance of convincing the opposition until a new scheme of economic theory has been developed and worked out. In the past we have been opposing the orthodox school more by our flair and instinct than because we had discovered in precisely what respects their theory was wrong.” But unfortunately, Keynes falls short in seeing that just like flair and instinct, theory is not sufficient in convincing the other side and for a good reason. Theory responds to need! By appropriating the Kuhnian quote “science progresses one funeral at a time” we would like to say that the favourite macroeconomic theory changes one crisis at a time but the underlying question remains the same and unanswered. We worry that the discourse among economists is controlled too much by the thought police of the “Top-five” journals which in the macroeconomics area insist on deploying versions of the

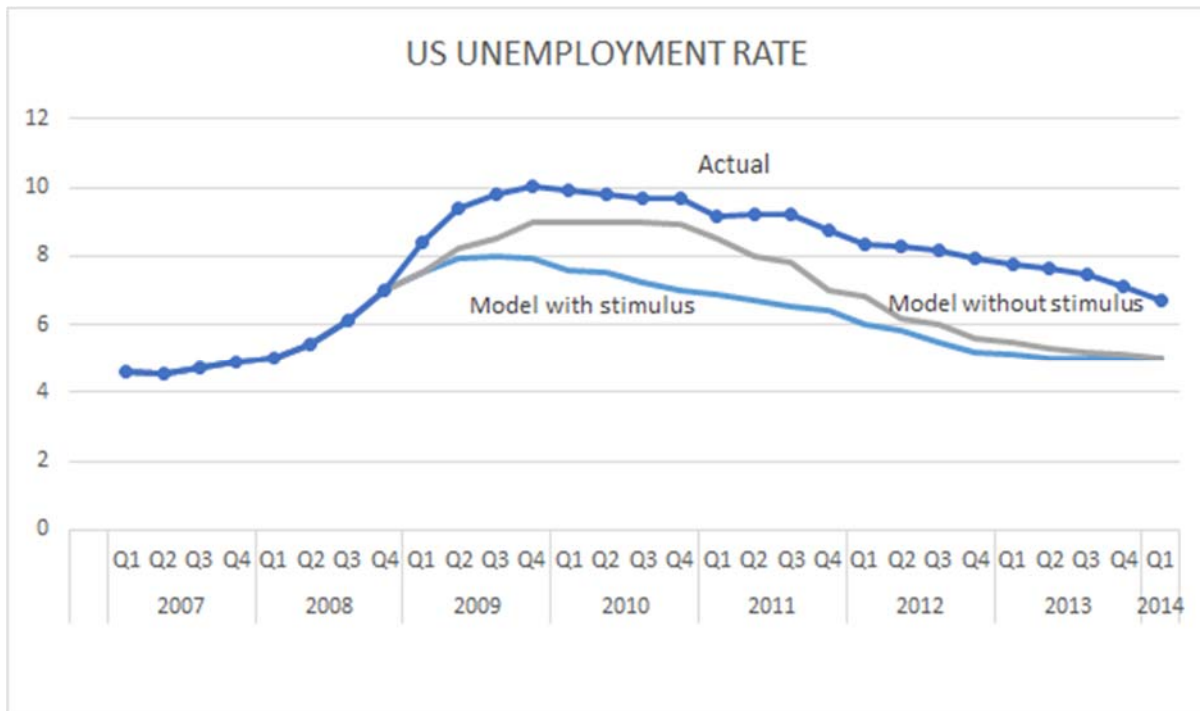
DSGE models, not admitting that this acronym captures three big things that economists do not understand: (1) How individuals organize their spending over time, (2) How they deal with statistical uncertainty and (3) How they are influenced by all the people around them.

5.2 Response to The Great Recession

The open letters showcased earlier were written in response to the American Recovery and Reinvestment Act of 2009 which at the time was being debated in the Congress. The act was an ambitious piece of legislation which ended up costing \$840 billion (CBO 2014). The Act had three broad objectives and each objective aligned well with a different type of fiscal instrument. A third of the expenditure would go towards giving tax relief to middle class and business investment. Another third would provide support to those who were hurt most by increasing transfer payments through food stamps and Medicaid. The remaining third aimed to cover the government consumption and investment purchases in energy, education and infrastructure which were the president's priorities (Grabell 2012). Significant portions of the Act had fungible expenditure provisions which were meant for state and local governments to use to avoid layoffs that might happen due to reduced revenue and balanced budget requirements⁸.

So did the fiscal stimulus work? Depends on who you ask and how they decided to measure it. Ignoring the cost effectiveness of the policy, some have asked whether or not the act met its key goal which was to "save or create at least 3 million jobs by the end of 2010" (Berstein and Romer (2009)). The non-partisan Congressional Budget Office estimated that in 2010 somewhere between 0.7 to 3.3 million jobs created or saved can be attributed to the Act. But numerous assumptions are required to draw empirical conclusions, and this case is no different. To arrive at these estimates, the CBO must create a counterfactual world were the stimulus was not implemented, which it does using models. This exercise of creating counterfactual world is a precarious one which is evident from the **graph presented below** which depicts the actual outcomes and two runs of the model, with and without the stimulus. The model predicted an unemployment rate 2 percentage points lower with the fiscal stimulus than without (Berstein and Romer (2009)). But the actual path the economy which was aided by an actual stimulus had a substantially larger unemployment rate than even the worst-case model scenario. Should we believe that the actual would have been even worse if the fiscal stimulus had not been implemented? Really? Why? Each tribe has a different answer.

⁸\$ 318 billion were spent in intergovernmental transfers to state (Carlino and Inman (2013)).



But what about the fiscal multipliers? Did the ARRA spending create more economic value than it cost? CBO estimates for the multipliers applying to government purchases, transfer payments to individuals, and lower- and middle-income tax cut multiplier are said to be between 0.5 to 2.5, 0.4 to 1.8 and 0.3 to 1.5 respectively. These suffer from both omitted variable bias and simultaneity. But certain randomization in implementation of individual provisions in the stimulus can help us identify exogenous geographical variation and provide us clear instances where the value of evidence is less disputable albeit for few provisions of the law. A good survey of this literature by Chodorow-Reich (2019) concludes that multiplier estimates are “in the range of 0.76 to 3.93, with a cross-study mean of 2.1 and median of 1.9”. But importantly, these are local fiscal multipliers. From early theoretical work, it appears that extrapolation from local fiscal multiplier to national multiplier shrinks the size of the multiplier (Ramey (2020)). There remains a great deal of uncertainty which multipliers exceed one, and when.

Finally, remember that fiscal policy can be used to invest in sectors and regions of the economy which can help grow the future economy despite the short-term multiplier being less than 1. A third of the Act focused on investments in energy, education and infrastructure to aid growth in the decades ahead. The debate about the size of the short-term fiscal multiplier fits well with the two-tribe reality, but can lead to faulty conclusions regarding the value of the heterogeneous features that were part of the stimulus.

So bottom line regarding the impact of ARRA 2009 is that the stimulus package “became a Rorschach test of amorphous programs, on which critics of all stripes could project their dissatisfaction with government and politics in general” (Grabell 2012). Despite the unclear empirical evidence, in 2014, when [leading economist were asked](#) whether they agree with the statement that “the benefits of the stimulus will end up exceeding its costs”, 67 % of those who answered either *agreed* or *strongly agreed*. Interestingly, when responses were factored in based on economists’ confidence, the proportion of economist either agreeing or strongly agreeing went up to 75 %.

If economists are being asked, [why not ask the common folk too](#)? In 2012, 37 % approved and 41 % disapproved. These approval ratings were strongly split along the party lines. Only 12 % Republicans approved where as 66 % of Democrats approved but 76 % of Republicans disapproved whereas only 16 % of Democrats disapproved. To say that fiscal policy is influenced by the ebb and flow of the politics might be an understatement. After the announcement of the fiscal stimulus on 22nd November 2008, it took 87 days of deliberation between Congress and the Executive branch to pass the ARRA 2009. Despite these 3 months of deliberation, the support for the Act was divided along party lines. Due to the Democrats' comfortable majority in the House they did not seek any Republican support. Hence no Republican congressmen voted in support of the Act. On the other hand, in order to avoid a minority filibuster, three Republican senators were persuaded to vote in favor of the bill.

Will it be the case that we generate unquestionable evidence such that a consensus is developed around efficacy of fiscal stimulus of 2009? Most probably no.

6. New direction of research

While the miasma theory and the contagion theory dominated European medical practice in the middle ages, the germ theory had adherents in Europe as well as the Middle East. It was the microscope first used by van Leeuwenhoek in the 1670s that contributed to the increasing dominance of the germ theory and also the experimental work initiated by Louis Pasteur in the 1860s. Though economists will never be able to conduct experiments to test the efficacy of fiscal policy, they can use the equivalent of a microscope by following the lives of the organisms that comprise the economy: the people, the families, the businesses. The Chetty et al. (2020) study of fiscal policy during the 2020 Covid-19 shutdown uniquely uses a microscope to draw conclusions regarding the lack of impact of fiscal policy on employment in the first few months of the downturn:

“We build a publicly available platform that tracks economic activity at a granular level in real time using anonymized data from private companies. We report daily statistics on consumer spending, business revenues, employment rates, and other key indicators disaggregated by county, industry, and income group.

...

These results suggest that traditional macroeconomic tools – stimulating aggregate demand or providing liquidity to businesses – may have diminished capacity to restore employment when consumer spending is constrained by health concerns.” Abstract

In other words, Chetty et. al. still think that fiscal policy works for normal recessions, but not for pandemic-induced shutdowns. Our view is that it was the fear of face-to-face contact that drove down employment in the face-to-face service sectors including restaurants and nail salons and dentist offices. When these workers lost their jobs but their incomes were made whole by fiscal policy, they did not use their incomes to visit restaurants or nail salons or dentist offices. This should not be thought to be a fiscal stimulus but a moral obligation. When the fear of Covid-19 abates, that is when a stimulus might accelerate the job formation in these hard-hit sectors. When Chetty et.al work with the data through 2021, they may find evidence of job-creation in response to fiscal policy, but maybe not.

The previous paragraph is a story that embodies the Chetty finding, which makes our final point – a deductive model is an essential part of a persuasive argument. Using a microscope to study the economy does not avoid the problems of drawing valid causal conclusions from non-experimental data. A microscope alone is not enough. It also takes skilled theorists to write compelling stories. The disease metaphor creates plenty of opportunities for theorists to create models that explain the growing fragility of the economy during expansions, the events that bring on the recession disease, the determinants of the seriousness of the recession, the medicine that works during the expansions and during the recessions, and what makes for a quick recovery. And these stories can be supported by data, though not confirmed. There is much yet to be learned about fiscal stimuli.

References

- Agarwal, S., Liu, C., and Souleles, N. S. 2007. "The reaction of consumer spending and debt to tax rebates evidence from consumer credit data". *Journal of Political Economy*, 115 (6), 9861019.
- Appelbaum, Binyamin. 2019. "The Economists' Hour: False Prophets, Free Markets, and the Fracture of Society"
- Auerbach, Alan J. and Yuriy Gorodnichenko, 2012. "Fiscal Multipliers in Recession and Expansion," NBER Chapters, in: *Fiscal Policy after the Financial Crisis*, pages 63-98, National Bureau of Economic Research, Inc.
- Barro, Robert J. and Charles J. Redlick, 2011. "Macroeconomic Effects From Government Purchases and Taxes" , *The Quarterly Journal of Economics*, Volume 126, Issue 1, February 2011, Pages 51102,
- Blanchard, Olivier, and Perotti Roberto (2002). "An Empirical Characterization of the Dynamic Effects of Changes in Government Spending and Taxes on Output." *Quarterly Journal of Economics*, 117, 13291368.
- Baily, Martin Neil, Robert E. Litan and Mathew S. Johnson, *The Origins of the Financial Crisis, Fixing Finance Series, Paper 3*, November 2008, Brookings.
- Bewley, Truman F. *Why Wages Don't Fall During a Recession* (1999) 1-19
- Bewley, Truman F. Comment on George A. Akerlof, William T. Dickens, George L. Perry, "Near-Rational Wage and Price Setting and the Long-Run Phillips Curve,;" *Brookings Papers on Economic Activity*, Vol. 2000, No. 1 (2000), pp. 1-60.
- Clemens, Jeffrey, and Stephen Miran. 2012. "Fiscal Policy Multipliers on Subnational Government Spending." *American Economic Journal: Economic Policy*, 4 (2): 46-68.
- Chetty, Raj, John N. Friedman, Nathaniel Hendren, Michael Stepner, and the Opportunity Insights Team, "How Did COVID-19 and Stabilization Policies Affect Spending and Employment? A New Real-Time Economic Tracker Based on Private Sector Data" Working paper, June 2020
- Chodorow-Reich, Gabriel. 2019. "Geographic Cross-Sectional Fiscal Spending Multipliers: What Have We Learned?" *American Economic Journal: Economic Policy*, 11 (2): 1-34.
- Duppe, Till, E. Ry Weintraub (2014) **Finding Equilibrium**, Princeton University Press, Princeton.
- Grabell, Michael (2012) "**Money Well Spent?** The Truth Behind the Trillion-Dollar Stimulus, the Biggest Economic Recovery Plan in History"
- Hansen, L.P. and James J. Heckman (1996) "The Empirical Foundations of Calibration," *Journal of Economic Perspectives*, Vol 10, No1, Winter, 87-104.
- Hawtrey, R. G. "Public Expenditure and the Demand for Labour." *Economica*, no. 13 (1925): 38-48. Accessed October 31, 2020. doi:10.2307/2548008.
- Head ML, Holman L, Lanfear R, Kahn AT, Jennions MD (2015) The Extent and Consequences of P-Hacking in Science. *PLoS Biol* 13(3): e1002106. <https://doi.org/10.1371/journal.pbio.1002106>

- Homan, P.T. 1941. Letter to Professor Paul A. Samuelson rejecting the paper with Stolper.
- Kahn, R. F. "The Relation of Home Investment to Unemployment", *The Economic Journal*, Volume 41, Issue 162, 1 June 1931, Pages 173–198, <https://doi.org/10.2307/2223697>
- Keightley, Mark P. (2019) "Fiscal Policy Considerations for the Next Recession", Congressional Research Service, R45780
- Leamer, Edward E. (1978) **Specification Searches: Ad Hoc Inferences With Non-experimental Data**, John Wiley and Sons, New York.
- Leamer, Edward E. , **The Craft of Economics Lessons from the Heckscher-Ohlin Framework**, Ohlin Lecture, M.I.T. Press, 2012.
- Leamer, Edward E. **Macroeconomic Patterns and Stories**, Springer, 2009.
- Leamer, Edward E. "Housing IS the Business Cycle," in Housing, Housing Finance and Monetary Policy, A Symposium Sponsored by the Federal Reserve of Kansas City, 2008. Revised and updated for the Encyclopedia of Finance.
- Leamer, Edward E. "Housing Really is the Business Cycle," *Journal of Money, Credit and Banking*, Supplement to Vol. 47, No. 1 (March–April 2015).
- Leamer, Edward E. "Why are Some Recoveries Rapid and Others Slow?", Anderson Business Forecast Report, June 2020
- Mill, John Stuart (1836), "On the Definition of Political Economy and the Method of Investigation Proper to It," Reprinted in **Essays on Some Unsettled Questions of Political Economy** (1844), 3d ed., London: Longmans Green & Co., 1877, pp. 120-164.
- Nakamura, Emi and Jon Steinsson, "Fiscal Stimulus in a Monetary Union: Evidence from US Regions", *American Economic Review*, 2014, 104 (3), 753-92.
- Parker, Jonathan A. 2011. "On Measuring the Effects of Fiscal Policy in Recessions." *Journal of Economic Literature*, 49 (3): 703-18.
- Romer, Christina D., and David H. Romer. 2010. "The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks." *American Economic Review*, 100 (3): 763-801.
- Samuelson, Paul A. (1947),. *Foundations of Economic Analysis*, Harvard University Press.
- Samuelson, Paul A. (1952), "Economic Theory and Mathematics – An Appraisal," *American Economic Review*, 42(2), pp. 56–66.
- Samuelson, Paul A., and Robert M. Solow. "Analytical Aspects of Anti-Inflation Policy." *The American Economic Review* 50, no. 2 (1960): 177-94. Accessed October 31, 2020. <http://www.jstor.org/stable/1815021>.
- Skidelsky, Robert. *Money and Government: The Past and Future of Economics*. New Haven; London: Yale University Press, 2018. Accessed October 31, 2020. <http://www.jstor.org/stable/j.ctv6gqq16>.
- Stolper, Wolfgang F., and Paul A. Samuelson. 1941. Protection and real wages. *Review of Economic Studies* 9 (4): 58-73