ECONOMIC RESEARCH
FEDERAL RESERVE BANK OF ST. LOUIS

WORKING PAPER SERIES

The Poverty of Macroeconomics --- What the Chemical Revolution Tells Us about Neoclassical Production Function

Authors	Yi Wen
Working Paper Number	2021-001A
Creation Date	February 2021
Citable Link	https://doi.org/10.20955/wp.2021.001
Suggested Citation	Wen, Y., 2021; The Poverty of Macroeconomics --- What the Chemical Revolution Tells Us about Neoclassical Production Function, Federal Reserve Bank of St. Louis Working Paper 2021-001. URL https://doi.org/10.20955/wp.2021.001

Federal Reserve Bank of St. Louis, Research Division, P.O. Box 442, St. Louis, MO 63166

The Poverty of Macroeconomics

—— What the Chemical Revolution Tells Us about Neoclassical Production Function

Yi Wen[1]

Abstract

Quantitative macroeconomics is often portrayed as a science—because of its intensive use of high-powered mathematics—with the possible limitation of being unable to conduct controlled experiments. To qualify as a science, however, theories in that discipline must meet a minimum number of criteria: (i) It has *explanatory* power to explain phenomena; (ii) it has *predictive* power to yield quantifiable and falsifiable statements about new phenomenon; and (iii) it has *operational* power to change the world.

A scientific theory consists of axioms and working hypotheses that facilitate the derivation of contestable statements from the axioms.[2] Hence, simply laying out a list of contradictions between a theory's implications and the data is often insufficient to disqualify a theory as science; it may have just challenged its working hypotheses, not its axioms. But, challenging a theory's working hypotheses is a crucial step to improve or falsify a theory. This is why Isaac Newton spent so much effort in his *Principia Mathematica* to deal with the law of motion under air friction.

This article discusses one of the working hypotheses of the Arrow-Debreu paradigm and its dynamic stochastic general equilibrium reincarnation in quantitative macroeconomics—the supply curve and its embodiment in the neoclassical production function. The supply curve is a much stronger pillar than the demand curve in holding up the Arrow-Debreu paradigm, but we argue in this article that the neoclassical production function embodying the supply curve is full of cracks.

More specifically, we show that the neoclassical production function is not quantifiable as a working hypothesis to support the Arrow-Debreu DSGE model, unlike the chemical reaction equations based on Lavoisier's oxygen theory of combustion. The neoclassical production function relies on the unobservable and unmeasurable Solow residual to explain the quantity of output produced at the firm, industry, or national level, and the hypothetical factors of production (capital and labor) are much like "fire, air, water, and earth" in the ancient Greek theory of the universe.

[1] Federal Reserve Bank of St. Louis. The views expressed are those of the individual authors and do not necessarily reflect the official positions of the Federal Reserve Bank of St. Louis, the Federal Reserve System, or the Board of Governors. I would like to thank Chun Xia and others for comments.

[2] For example, Galileo's law of free fall states that, in the absence of air resistance, all bodies **fall** with the same acceleration, independent of their mass. It consists of the axiom of mass-independent acceleration and the working hypothesis of a vacuum (or the lack of air friction). To qualify for a science, however, either the axiom of constant acceleration or the working hypothesis of air friction must be itself quantifiable. This is why simply observing that two pieces of wood fall at different speeds through the air does not falsify Galileo's theory because the theory's working hypothesis (or working condition) is not met in the experiment. However, once the exact degree of friction is quantitatively measured, Galileo's theory should also yield precise predictions in the modified environment. In other words, the theory of acceleration can also be proven under the condition of air friction as long as the exact degree of friction is quantifiable. Newton's aether theory of the universe was abandoned because neither the axiom nor its working hypothesis is quantifiable.

Because the working hypotheses of quantitative macroeconomics are not themselves quantifiable, the neoclassical theory is not yet a science. And this explains the lack of power for DSGE models to predict the 2008 Financial Crisis and the inability of economic theory to change the world by engineering or recreating economic prosperity in developing countries.

1. The Arrow-Debreu paradigm

Neoclassical economics is often portrayed as a science—because of its intensive use of high-powered mathematics, except with the possible limitation in conducting controlled experiments. Here is Robbins' famous all-encompassing definition of **economics** that is still used to define the subject today: "**Economics** is the **science** which studies human behavior as a relationship between given ends and scarce means which have alternative uses." [3]

To qualify as a science, however, a theory must meet a minimum number of criteria, such as: (i) It has "explanatory power" to explain a set of phenomena or facts of interest to its practitioners; (ii) it has "predictive power" to yield quantifiable and falsifiable statements or counterfactuals; and (iii) it has "operational power" to change the world, such as suggesting ways to not only recreate the phenomenon the theory was designed to explain but also create new phenomena that do not yet exist in nature—because "knowledge is power."

For example, chemistry is a science. Within this branch of knowledge, a chemist can (i) explain the phenomenon of combustion, such as gunpowder explosions; (ii) predict that the intensity of combustion depends on the amount of oxygen, and oxygen itself is measurable and quantifiable; and (iii) help scientists to design or create new combustible materials (such as the nitrogen bomb) based on the oxygen theory. Among the three criteria, the last one is perhaps the most important and crucial to determine whether a theory or branch of knowledge is a science or a pseudo-science. However, the operational power of a theory often lies in the quantitative measurability of its working hypotheses (or working conditions).

For instance, the phlogiston theory dominated the fields of both chemistry and natural philosophy during the 17th to 18th centuries, before Lavoisier's Chemical Revolution. It offered reasonable explanations for combustion; and it had predictive power that combustion could not take place in the absence of "air." But, neither phlogiston nor its working hypothesis is measurable or quantifiable, and as a consequence the theory has no operational power to create new combustible materials.

The ability to possess operational power is important. When a theory is incorrect or is pseudo-scientific, it may appear to have explanatory power, because people can always choose to ignore contradictions. However, when a theory is pseudo-science, it is useless in generating practical knowledge, hence violating the Baconian principle that "knowledge is power." The Baconian notion of "power" simply means "causal effects": True knowledge about nature must consist of causal statements in the sense that removing or presenting the causes should eliminate or recreate the phenomenon the theory intends to explain under its working hypotheses.

Sometimes a pseudo-science may continue to be accepted and held by the research community because there is not yet a better theory to replace it. Sometimes a pseudo-science may even continue to be believed after a true scientific theory has already emerged. For example, the phlogiston theory continued to be accepted by many famous chemists, physicists, and mathematicians in the scientific community even after Lavoisier forcefully demonstrated that combustion was caused by the presence of oxygen and that oxygen can be quantitatively measured and recreated from water or other chemical compounds. The scientists of the day found it difficult to accept Lavoisier's new theory because oxygen

[3] See https://www.econlib.org/library/Enc/bios/Robbins.html.

at the atomic level was too small to be directly observable by the technology at hand. But the power of science precisely lies in its explanatory power, predictive power, and operational power based on measurability, not necessarily on observability. Seeing can be deceiving. Also, horoscopes and the biblical interpretation of the origin of life are still widely believed today. But, such pseudo-sciences or religious beliefs clearly lack "operational power" in recreating phenomena and new facts because the working hypotheses of these theories are not quantifiable or measurable.

Modern macroeconomic theory builds on the Arrow-Debreu model as its micro-foundation. The building blocks (axioms) of the Arrow-Debreu model include (i) the demand function, (ii) the supply function, and (iii) a Walrasian market-clearing mechanism under which a set of equilibrium prices can emerge to equate the demand and supply in each commodity market.

By applying the fixed-point theorem in mathematics (calculus), economists can prove analytically that under certain regularity conditions (working hypotheses) imposed on the demand and supply functions, a set of market-clearing equilibrium prices always exist and are often unique. For this reason, the Arrow-Debreu model is also called the theory of value or "price theory." The equilibrium outcome corresponding to the market-clearing prices is called an equilibrium allocation of scarce resources, featuring the equality between the quantity of supply (implied by profit-maximizing firms) and the quantity of demand (implied by utility-maximizing consumers).

Under a set of additional working hypotheses (market conditions)—such as that markets are complete and perfectly competitive, information is public and fully transparent, agents are rational and pursue nothing but their own self-interests, and there are no public good/bad and externalities in people's actions—a striking implication of the Arrow-Debreu model emerges: the two Welfare Theorems.

The First Welfare Theorem states that any competitive equilibrium (or Walrasian equilibrium) leads to a Pareto-efficient (welfare-maximizing) allocation in scarce resources.[4] The Second Welfare Theorem states the converse, that any efficient (welfare-maximizing) allocation can be sustained in a competitive equilibrium.[5] So, given that economic efficiency means that social welfare is maximized, a competitive market economy based on price mechanisms and profit-seeking motives is the most efficient in resource allocations.

The welfare theorems are often taken to be analytical (mathematical) proofs of Adam Smith's "invisible hand" hypothesis and a scientific support for free-market capitalism and the non-interventionism ideology: Let the markets do the work and the outcome will be efficient, and there is no need for government to intervene under any circumstances (except praying for God to give us good weather or good technologies such as fire or gunpowder). Or as President Ronald Reagan famously said: "The government does not solve the trouble; it is the trouble." [6]

---

[4] A Pareto-efficient allocation means that the resource allocation cannot be further improved without making someone in the economy worse off. However, there may exist multiple such allocations and none of them is universally "desirable" by all agents in the economy.

[5] For simple reference, see http://en.wikipedia.org/wiki/Fundamental_theorems_of_welfare_economics. For sophisticated readers, see Mas-Colell et al. (1995).

[6] Notice that the welfare theorems can also be taken as confirmation or support of central planning economies, provided that the government is altruistic or benevolent and has perfect information (as the agents do) on the economy.

Neoclassical macroeconomics builds on the Arrow-Debreu model by introducing the aggregate production function, the aggregate consumption function, the aggregate capital-investment function, and a full set of Arrow-Debreu Securities under aggregate uncertainty into the Arrow-Debreu framework. Such a macroeconomic model is often called a dynamic stochastic general equilibrium (DSGE) model, because demand and supply decisions are derived from dynamic optimization in a stochastic environment.

More specifically, a DSGE model has several building blocks consistent with the Arrow-Debreu framework, namely, a set of well-behaved intertemporal demand and supply functions in (i) the goods markets, (ii) the factors markets (such as labor, capital, and intermediate goods), and (iii) the asset markets (such as money, government bonds, and private-issued securities), which are all derived from dynamic optimization problems under uncertainty where the underlying distributions of the exogenous shocks are known to firms and consumers.

Under the same set of working hypotheses regarding (i) mathematical properties of utility/production functions (that derive the demand/supply functions) and (ii) the market structures as in the Arrow-Debreu model (such as instantaneous market clearing via the Walrasian process and perfect competition), the two Welfare Theorems also hold in a DSGE model.

Besides the two welfare theorems, there are at least six additional predictions (implications) that can be deduced from a DSGE model:

1. Say's Law holds; namely, aggregate supply determines aggregate demand. Since anything produced in the economy will be redistributed in terms of payments (income shares) to different parties involved in the production process, aggregate demand will exactly exhaust aggregate supply at the market-clearing prices. There will not exist any unutilized resources (such as unemployed labor) or excess (idle) production capacities.

2. Money is neutral and monetary policies have no effect on the economy's resource allocation.

3. When production technologies are public knowledge or publicly available, there should be no difference between an agrarian economy and an industrial economy in terms of living standards. Under free trade, countries will automatically specialize according to their comparative advantages and the equilibrium terms of trade, overtime income per capita will converge to the same level across all countries even if they started with different levels of natural endowments and capital.

4. There is no need for any institutions to exist, such as the nation state, the legal system, financial regulations, central banks, industrial policies, let alone any ideologies.

5. Long-run economic growth is possible only through productivity growth, and short-run economic fluctuations are caused only by random changes in technologies, peoples' preferences, or government policies such as tax rates. So it is impossible for the Great Depression and the 2008 Financial Crisis to occur unless the economy's production technology suddenly breaks down, or people suddenly decide to enjoy more leisure and refuse to work, or the government opts to increase taxes on the private sector.

6. The free market will react to exogenous shocks optimally, and any interventional government policies designed to counter the adverse effects of the shocks are doomed to be counterproductive (except in the case where the shocks are from the government itself). Industrial policies serve only to distort

economic incentives rather than promote economic development and growth and, hence, must be abandoned.

Some of these implications are obviously false. For example, the Great Depression was certainly not caused by changes in weather, people's taste for leisure, or the tax rates. But this by itself does not constitute a refutation to the Arrow-Debreu model. Remember that Newtonian mechanics predict that all free-fall bodies starting at the same height will reach the ground at the same time regardless of their shapes and masses? The fact that in reality this does not happen does not by itself constitute a refutation to the theory, because of the existence of air friction.

Therefore, economists believe that after the establishment of the Arrow-Debreu framework, the only job left is to find the right type of "air friction" to explain the real-world economic phenomena. For example, if price adjustment is costly, then money is no longer neutral. If households are borrowing constrained, then financial institutions matter. If people are not rational or information is not perfect, then the 2008 Financial Crisis or the Great Depression could be explained by miscalculations of firms or households.

Hence, simply laying out a list of contradictions between theory and data seems insufficient to disqualify a theory as scientific; it may have only falsified the working hypotheses (regularity conditions) in the theory. The basic axioms in the theory may still be correct and provide a unifying principle to explain the world. So if the working conditions are relaxed or made more realistic, then the quantitative predictions of the theory may explain the real world better or match the data more accurately.

Therefore, it seems important to differentiate the fundamental axioms of a theory from its working hypotheses under which qualitative and quantitative predictions are derived. In Newtonian mechanics, for example, one of the fundamental axioms (assumptions) is the gravity principle, and one of the working hypotheses to derive its predictions is a vacuum or the absence of air friction.

However, one must note this: Either the concept of gravity or the working hypothesis—the so-called "air friction"—must itself be quantitatively measurable and falsifiable. Otherwise Newtonian mechanics is not scientific, but a religious belief similar to the four elements theory of the ancient Greeks. This was exactly why Newton's aether theory of the universe was abandoned, because neither the concept of aether nor its working condition is measurable and quantifiable.[7] This is also why Galileo used iron balls instead of two pieces of wood in his experiments to prove his concept of uniform acceleration under gravity, because air friction is negligible in resisting iron balls and the degree of air friction is potentially quantitatively measurable in experiments.[8]

Analogously, the fundamental axioms of neoclassical economics are the demand-supply and Walrasian market-clearing principles in the Arrow-Debreu model. And the set of regularity conditions needed to derive the Welfare Theorems, such as the mathematical properties of the demand-supply functions, the

[7] In Newtonian physics, **aether theory** (also known as **ether theory**) proposes the existence of a medium, a space-filling substance or field, thought to be necessary as a transmission medium for the propagation of light and electromagnetic or gravitational forces. Since the development of the theory of special relativity by Einstein, the aether theory was abandoned by modern physics.

[8] In fact, Newton himself devoted much effort in his *Principia Mathematica* to study the law of motion of cannon balls under air friction.

completeness of markets, the full rationality of self-interested agents, the perfection of information and so on, are the theory's working hypotheses.

For a theory to be scientific, however, either its axioms or its working hypotheses must be quantifiable and refutable. A theory cannot be qualified as a science if the hypothetical working conditions to support the axioms are not quantifiable and refutable. For example, for the Christian belief of God to be a science, the working conditions to reveal or predict God's will or love must be quantifiable and measurable.

How does the situation look for neoclassical economics or DSGE models? For the sake of argument, I will focus in this article on the supply curve of the neoclassical theory—specifically, on the measurability of the production function used in DSGE models. My point is to show that if the supply side is already troublesome, the demand side can only be worse.

As mentioned above, if the working hypotheses of a theory (such as air friction in Newtonian mechanics) are not themselves quantitatively measurable and are constantly contradicted by facts, then the theory as a whole is not yet a science, or it may just be a pseudo-science. This was exactly what happened to the phlogiston theory of combustion in the history of chemistry, because chemists and natural philosophers could not find quantitatively measurable working hypotheses to support this theory, which is also the case for the concept of God.

Most importantly, the lack of "operational power" in a theory is often the direct consequence of the failure (or misspecification) of the working hypotheses that are keys for a theory to work. This is why a necessary and crucial step in improving or refuting a theory is to quantitatively evaluate or falsify its working hypotheses rather than directly attack its axioms.

2. Trouble with the neoclassical production function

Among the six implications of a DSGE model listed above, Say's law is the most far-reaching and fundamental. Say's law holds under a set of working hypotheses, and the most important working condition for Say's law to yield quantitative predictions in any DSGE model is the aggregate production function—it hides the secrets of not only long-run economic growth but also short-run economic fluctuations.

Everything else in a DSGE model, such as the consumption function and asset demand function, is more or less a side show, as they determine only how the aggregate output is redistributed across space and time in general equilibrium. In other words, under Say's law the size of the economic pie at each point in time is determined completely by the production function, and the rest of the model's setups determine how the pie is divided among different parties according to preferences and risk-sharing possibilities (such as consumption today versus consumption tomorrow and borrowing versus lending). The dynamic part of the model involves saving, investment, and intertemporal lending. An equilibrium steady state (or balanced growth path) can be reached in finite time such as where the desired level of investment (savings) equals the natural depreciation rate of the capital stock such that the aggregate capital stock becomes constant (or grows at constant rate) in the long run. To simplify the picture, I will ignore the dynamics and focus on the steady state where the capital stock is constant. My argument can be easily generated to a dynamic setting where the capital stock is not a constant.

More specifically, the DSGE model assumes that there exists an aggregate production function that relates inputs to output:

$$Y = AF(K, N), \tag{1}$$

where *Y* is output; *F* is a mathematical mapping or function that relates capital *K* and labor *N* to output; and the coefficient *A* is called the total factor productivity (TFP) or simply the Solow residual, which is exogenous and can have both a long-run growth trend and a short-run stochastic component. Since *A* Is exogenous, the stochastic component in *A* (or random changes in TFP) is also called "technology shocks."

A convenient functional form of *F* is specified as a Cobb-Douglas function:

$$Y = AK^{\alpha}N^{1-\alpha}; \tag{2}$$

where $\alpha \in (0,1)$ is called the output elasticity of capital and $(1-\alpha) \in (0,1)$ is called the output elasticity of labor. The values of these elasticities imply that (i) total output is increasing in both capital and labor at a diminishing rate and (ii) the production technology exhibits constant returns to scale—implying that, given A, total output will be doubled when both inputs are doubled. Obviously, the higher the TFP the more output is produced for a given amount of inputs. Although the growth in the capital stock and labor force can also cause long-run growth in income *Y*, Solow (1956) proved that under the assumption of constant returns to scale for the production function, per capital income or the output-to-labor ratio $\frac{Y}{N}$ in the long run is driven only by TFP because of diminishing returns to capital ($\alpha < 1$): namely, the capital-to-labor ratio $\frac{K}{N}$ will be counterbalanced by the force of capital depreciation such that it will converge to a constant in the long run, leaving TFP (A) as the only possible force driving economic growth in per capita output in the long run.

The TFP level *A*, however, is not observable from the data, but can be backed out by utilizing the production function and data on {*Y, K, N*} from the National Income Accounts and Government Statistical Bureau, so the Solow residual can be indirectly "measured" by inverting the production function:

$$A = \frac{Y}{F(K,N)} = \frac{Y}{K^{\alpha}N^{1-\alpha}}, \tag{3}$$

where *Y* is now replaced by the actual gross domestic product (GDP) in a nation, *K* is the "actual" aggregate capital stock (often derived as the sum of the market values of past aggregate investment), and *N* is the actual level of aggregate employment.

Notice the "trick" here: "*A*" is assumed to be an independent exogenous force driving long-run output growth, but it is unobservable. So, macroeconomists propose to use the actual output *Y* to indirectly measure *A* using the hypothetical production function *F*. This implies, however, that even if the production function is completely misspecified or simply false, the measured Solow residual according to the above equation always exists, and there is no way to know whether the Solow residual so measured is correct or not. In other words, any misspecifications of the production function *F* or measurement errors in *K* and *N* will be captured by Solow residual *A*, but we have no idea how *A* should behave in the first place.

Taking derivatives with respect to time on both sides of the production function gives the percentage changes in TFP as the difference between the changes in total output *Y* and the changes in total inputs:

$$\Delta\%A = \Delta\%Y - \Delta\%F(K,N) \tag{4}$$

$$= \Delta\%Y - \alpha\Delta\%K - (1-\alpha)\Delta\%N.$$

This is why "*A*" is called the Solow residual, as Robert Solow (1956) first adopted this aggregate production function to study economic growth and this technology coefficient *A* accounts for the gap (residual) between the actual output *Y* in the real world and the model-implied output *F(•)* and.

Consequently, if we measure TFP this way, then even if the production function is completely false and the so-called aggregate capital never exists or is taken from the Moon, the model will still enable us to predict the movements in aggregate output Y perfectly once we are given the measured Solow residual *A*.

And this circular reasoning or tautological exercise is exactly what quantitative macroeconomists do all the time in their quantitative research, ever since the publication of Kydland and Prescott (1982). Namely, they assign to the Solow residual in their DSGE model the same statistical properties as the estimated Solow residual from the data and then feed the "calibrated" Solow residual back to their original models to "explain" or predict output *Y*.

Of course, a typical DSGE model often involves more variables than output or capital. It also involves, for example, consumption *C*, investment *I*, asset prices *P*, and so on. But, a DSGE model typically has the same number of equations (derived from utility/profit maximization) as the number of choice variables. Namely, for each choice variable, such as {*Y, K, N, C*}, there is a corresponding equation called the decision rule that relates an endogenous variable to the rest of the economy.

For example, if a DSGE model hypothesizes the existence of a representative household that chooses aggregate consumption *C* to maximize a utility function *U(C)* subject to a budget constraint or aggregate resource constraint, then the optimization process will yield a decision rule for consumption in the form of *C(Y,X)* that relates consumption C to aggregate income *Y* and a set of other variables *X* in the model.

The consumption function *C(Y,X)* in principle should be used to test the model by comparing it with the actual consumption $\hat{C}$ in the data. But no DSGE model can pass the test since nobody knows how people actually behave and what type of utility functions they have (if there are any such utility functions at all). But this does not matter to quantitative macroeconomists because they can always introduce another shock called "consumption shock" β to fill in the gap between the model-implied consumption and the real-world consumption data $\hat{C}$ in the following way:

$$\hat{C} = \beta C(Y,X); \tag{5}$$

where the consumption shock $\beta = \frac{\hat{C}}{C(Y,X)}$ captures any gap between the model-implied consumption C(Y,X) and its empirical counterpart $\hat{C}$ in the real world. So a DSGE model can also match the consumption data perfectly because the imaginary shock "β" serves the exact role of the Solow residual *A* in the production function.

This is why a DSGE model needs many "imaginary shocks" in its quantitative exercises—in fact, as many shocks as the number of endogenous variables in the model.[9] For example, a 10-variable DSGE model requires 10 "shocks" to match the real-world data sets perfectly. This is why these shocks are also called "wedges."[10] These "wedges" are in general highly serially correlated over time and mutually correlated with each other. This explains the "success" of DSGE models in "explaining" aggregate economic fluctuations observed in any nation in any century on any planet in the universe.

One would therefore be tempted to conclude logically that if such "wedges" exist in a DSGE model and are very large in sizes (variance and covariance), they should be themselves hardcore evidence to reject the model, because they simply reflect the bad fit of theory to data. In other words, these wedges such as the Solow residual are nothing but "measures of our ignorance."[11]

However, macroeconomists do not think this way. Instead, they believe that their models are scientifically sound and such "wedges" are *real forces* driving the real-world economic fluctuations instead of imaginary forces created by DSGE model builders. In other words, they do not question the possible misspecifications of their models, such as the production function, utility function, or market structures. Instead, they think the "wedges" capture the *imperfections of the real world*.[12] They believe that their Arrow-Debreu model is the perfect world against which any economy on any planet should be benchmarked, as Greenwich Time (GMT) is the benchmark for all clocks in the world. So, from this perspective, the emergence of wedges from a DSGE model is simply because the real world itself is imperfect.

This is why macroeconomists never use the existence of "wedges" as a reason to reject or falsify their models; instead they use the "wedges" to provide "stories" to show how successful their models are in revealing the imperfections of the real world and in "explaining" the real world fluctuations. In other words, these wedges capture the gap between the real world and the ideal Platonic world that only macroeconomists can help find by using their scientific training in DSGE models. Thus, macroeconomists claim that they are contributing to our understanding of the real world by "discovering" these wedges.

Therefore, no matter how deadly wrong a DSGE model is in describing the real world, it is not refutable because macroeconomists can always find enough wedges such that the model matches the data exactly; and they can also claim at the same time that the real world is "proven" to be imperfect, simply because it does not agree with their Platonic model.

Let's pursue this reasoning a bit further. For the sake of argument, consider the extreme case where the DSGE model is deliberately designed in such a ridiculous way that it is completely uninformative and deadly wrong, say, by assuming that the production function *F(K,N)* and the consumption function *C(Y,X)* are both constant and equal to 1, no matter how many inputs are used to generate output and utilities. Namely, any amount of capital and labor can yield one and only one unit of output. Then, via the same exercise outlined above, the model builder would find that the Solow residual would be identical to the

[9] Romer (2016) coined the term "imaginary shock."
[10] See Chari, Kehoe, and McGrattan (2007).
[11] Abromovitz (1956) famously referred to the Solow residual as "the measure of our ignorance."
[12] See Chari, Kehoe, and McGrattan (2007).

real-world GDP: $A = Y$, and similarly the consumption wedge would be identical to the real-world consumption: $\beta = \hat{C}$. This extreme model is clearly an "empty shell" model by our design because there is nothing in it.

Yet, even in this ironic scenario, macroeconomists can still claim success in "showing" that the real world is imperfect and the wedges are driving the real world fluctuations. They can still invent stories behind the movements of the wedges to sound like empirical economists. For example, the Solow residual A and the consumption wedge β so derived would certainly capture the Great Depression and the 2008 Financial Crisis when U.S. GDP and actual consumption data are used in backing out the wedges. In this case, the model builder would blame the Great Depression on a highly persistent negative technology shock and an adverse consumption shock. Fancy econometrics can also be performed to reveal the respective contributions of the two wedges to the variance of GDP and consumption. The researcher will find that TFP shocks explain 100% of U.S. GDP and x% of U.S. consumption and that consumption shocks in turn explain 100% of *C* and z% of GDP. Nice! The researcher can then invent stories by citing tabloid newspapers that, during the Great Depression, U.S. technologies were stolen by the Russians and the American consumers were so upset they reduced consumption sharply.

Since there is no way to know whether the "wedges" are real or imaginary, no matter how the economic model is misspecified or deadly wrong, one can always interpret the "wedges" in any way at will. In general, among all such imaginary wedges found in any DSGE model, economists often find that the TFP shock "*A*" is far more important than any other wedges in explaining the movements in output *Y* and employment *N*, while preference shocks and other shocks appear more "relevant" in explaining consumption C and asset prices.[13] Since this stylized fact is somewhat "unexpected," it has even boosted confidence for economists.

The reason is actually simple. In an Arrow-Debreu framework, aggregate supply determines aggregate demand. Although on the surface the aggregate demand is determined by households' utility maximization subject to household budget constraints and conditions in the financial markets where savings are determined, it is ultimately pinned down by aggregate supply under Say's law. According to Say's law, anything produced by the production side will be distributed in terms of wages, profits, and tax revenues to different parties. Therefore, there cannot possibly exist excess supply or excess demand, so whatever is produced will be demanded. This is why ultimately, in general equilibrium, the key factor that can move the production frontier and the size of economic pie around is the productivity shock A. If TFP increases over time, so do aggregate supply and aggregate demand. If TFP moves up and down, so do aggregate supply and aggregate demand. However, since the pie is divided between consumption and investment in every period, we also need consumption and investment shocks to fill in the gaps between the model-implied consumption/investment and the actual real-world consumption/investment.

Consequently, a standard macroeconomic model would (i) explain economic trends mostly by the growth trend in TFP and (ii) explain recessions and booms mostly by the random components in TFP. The rest of the wedges simply determine how the aggregate pie is divided among different income earners, such as wages for workers (consumption) and profits for firms (investment).

---

[13] For example, see Smets, F., & Wouters, R (2007). Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach. *American Economic Review*, 93, 586-606.

Because a DSGE model can "match" the real-world data so well by allowing even just one shock—the measured Solow residual or TFP shock, Kydland and Prescott's seminal paper published in 1982 (earning them the Nobel Prize in 2004) asserts that they have "proven" that an Arrow-Debreu general equilibrium model provides an almost perfect description of the postwar U.S. economic fluctuations; which enables its followers to also conclude that there is little room or scope for government policies because most fluctuations were found to be caused by random movements in the technology, which by assumption is "exogenous" and cannot be affected by government policies.

Their argument is as follows: In times of bad weather, the harvest diminishes for any given level of labor and fertilizer inputs. So farmers' optimal response is to stay home by cutting back hours worked. But if the government uses stimulating policies to incentivize farmers to work harder during cold, rainy days, it would make people worse off because the market forces have already responded in an efficient way to the bad weather. On the other hand, when the weather is good, farmers opt to work harder and it amplifies the harvest brought about by sunshine and the right temperature. It would be counterproductive for the government to try to "cool down" the economy by increasing taxes because this "boom" is driven by good weather and it is better to let farmers work harder in sunshine than in rain. This is why either during the Great Depression or during the so-called "housing bubble" period of 2001-2007, the government should have left the economy alone. Any countercyclical policies are bad for people's welfare.[14]

No wonder the same DSGE model can provide "explanations" for any country's long-run economic performance and short-run fluctuations, making the DSGE framework a universal science in economics. For example, a well-cited article written by Hayashi and Prescott (2002) uses a DSGE model to examine the poor performance of the Japanese economy in the 1990s, which is a well-known decade of economic stagnation after the burst of the Japanese real estate bubble. They claim to find that the problem is not a breakdown of the financial system or the consequence of the burst of the Japanese real estate bubble, but rather the consequence of a period of low productivity growth in the Solow residual. They conclude in that paper that "growth theory, treating TFP as exogenous, accounts well for the Japanese lost decade of growth."[15]

This is why for virtually all economic booms and recessions across all countries, regardless of their economic structures and stage of development, their political systems, and their legal and financial institutions, macroeconomists always find that a technology shock is the most important driver of both long-run economic performance and short-run fluctuations in national GDP, whether that be the African countries in a poverty trap, the Latin America countries in a deep financial crisis, or the Asian countries in a postwar rapid growth phase.

Not surprisingly, a paper by Cheremukhin, Golosov, Guriev, and Tsyvinski (2015) showed that a DSGE model with random movements in the Solow residuals can very well "explain" China's economic performance since the founding of the People's Republic of China in the early 1950s all the way to the present time. And yes, they use the same DSGE model with the same production function and utility function, as if Chinese firms have been identical to American firms and have never upgraded their production functions and Chinese consumers have been identical to American consumers and have

[14] See Cole and Ohanian (1999, 2004)

[15] Hayashi, Fumio, and Edward C. Prescott. "The 1990s in Japan: A lost decade." *Review of Economic Dynamics* 5.1 (2002): 206-235.

never changed their perception of commercialism, because any such differences or changes (if exist) would be captured by the country-specific “wedges” anyway.

To make their story look more credible, these authors added a second production sector called “agriculture” into a standard DSGE model alongside the industrial sector; thus they introduced one additional Solow residual in the production function for food in addition to industrial goods. This additional wedge would of course fill in the gap between the model-predicted agricultural production and the actual agricultural production in rural China. This may reduce the importance of the industrial sector’s TFP shocks during the period from 1953 to 1978. But after the economic reform in 1978, since China experienced rapid industrialization and the agricultural sector shrunk rapidly, the industrial sector’s Solow residual starts to be more and more important in explaining China’s aggregate output and provide more sensible “interpretations” behind China’s bumpy transition from an agrarian society to an industrial society.

So, instead of interpreting the enormously larger Solow residuals as big failures of their model in describing China, they view them as big successes of the neoclassical paradigm in “explaining” the China miracle. By looking at those sectoral TFP shock processes, these authors “invent” stories about China’s policy failures during episodes of economic downturns, such as the Great Leap Forward movement, by showing a big drop in the measured Solow residuals in both sectors in 1958. But these authors did not really put such casually discussed policy failures literally into their model and show exactly how each policy of the Chinese government’s Five-Year Plan could have caused the fluctuations in China’s TFP.

Although for scientists or many non-economists and lay people, such an economic approach sounds ridiculous, yet it has been the dominating paradigm taught in U.S. graduate schools and top PhD programs all over the world since the early 1980s. In the past half century, many smart young people in developing countries (who are interested in learning economics to help their home country’s economic development and industrialization) have been trained in the U.S. by such economic theories. Those who are very good at mathematics eventually find jobs in top U.S. economics departments and become professors. The ones who aren’t so good at solving such dynamic macroeconomic models find jobs at the World Bank or the IMF.[16]

Is such a macroeconomic research paradigm really scientific? One critical criterion of science or scientific theory is that it has “operational power” to change the world. Unfortunately, macroeconomists have not been able to use DSGE models to recreate economic booms or industrial revolutions in any developing countries. Ironically, some people argue that the Chinese economic miracle took place precisely because the Chinese government refused to follow the prescriptions of orthodox economic theories, unlike the Russian government.[17]

Even the new institutional economists such as Douglas North and Daren Acemoglu may advocate that the “wedges” reflect the imperfections of the institutions in the real world, not the imperfections of the neoclassical model. In fact, they implicitly acknowledge that the real world would indeed behave like the one in the neoclassical model if institutions in the real world were perfect. They seem to also believe

[16] For a controversial statement about this fact, see Joseph Stiglitz, http://www.whirledbank.org/ourwords/stiglitz.html.

[17] See, e.g., Wen, Y. (2015). Lin, J. Y. (2012). Ruan, J., & Zhang, X. (2009)

that once the political institutions in developing nations become like U.S. institutions, these economies will be able to grow as fast as the U.S. and be as rich as the U.S. today.[18]

Is this really so? So far, only China has been able to grow at nearly a double-digit rate over a span of 40 years, yet China has not adopted the institutions or political systems of the U.S.

What these new institutional economists do not say is that even if the political institutions in all developing countries become identical to the American ones, their Solow residuals will not disappear and they would still need such "imaginary shocks" to explain their Great Depression and their financial crisis, as already demonstrated by Kydland and Prescott (1982) and Chari, Kehoe, and McGrattan (2002, 2007).

3. What the phlogiston theory tells us about quantitative macroeconomics

The problems of quantitative macroeconomics have not gone unnoticed, not surprisingly.[19] Despite such criticisms, macroeconomists have decided to ignore them, or have simply added more frictions and shocks into the DSGE model while keeping the "spirit." According to the 2017 Nobel laureate in economics Paul Romer (2016), this type of macroeconomics is a pseudo-science:

"*Macroeconomists got comfortable with the idea that fluctuations in macroeconomic aggregates are caused by imaginary shocks, instead of actions that people take, after Kydland and Prescott (1982) launched the real business cycle (RBC) mode*l. *The real business cycle model explains recessions as exogenous decreases in phlogiston*."[20]

Following Romer's analogy of the Solow residual as phlogiston, let us look deeper into the phlogiston theory in the history of chemistry, so as to use the Chemical Revolution that rejected the phlogiston theory to shed light on the troubles of neoclassical economics.

Before the Chemical Revolution that established chemistry as a science, alchemists and natural philosophers in the 18th century widely accepted the phlogiston theory as an explanation to the phenomenon of combustion as well as the phenomenon of respiration of animals and humans. The reason was based on the observation that air is critical for both combustion and respiration.

The phlogiston theory postulated the existence of a fire-like element called phlogiston, which is contained within combustible materials and released during combustion. The name comes from the

---

[18] See, e.g., Douglass C. North (1990), Institutions, Institutional Change and Economic Performance. Cambridge: Cambridge University Press, 1990; Acemoglu, D., & Robinson, J. A. (2012). *Why nations fail: The origins of power, prosperity, and poverty*. Currency.

[19] For the earlier debate in 1950s on neoclassical growth models and the issue of production function, see the Cambridge capital controversy at https://en.wikipedia.org/wiki/Cambridge_capital_controversy. For more recent concerns or criticisms on the neoclassical model, the real business cycle (RBC) model, or the DSGE model, see for example, Krugman (2009), Romer (2016), Blanchard (2016), Gali (2017), Stiglitz (2018), Xu (2020), the symposium on "Macroeconomics after the Financial Crisis," in the Fall 2010 issue of the *Journal of Economic Perspectives,* as well as the online discussions at https://www.econjobrumors.com/topic/what-is-a-negative-technology-shock/page/4.

[20] Romer (2016).

ancient Greek “φλογιστόν phlogistón (burning up) and φλόξ phlóx (flame). The idea was first proposed in 1667 (around the time Newton published his *Principia Mathematica*) by a German doctor and alchemist Johann Joachim Becher and later put together more formally by his student Georg Ernst Stahl in 1703. [21]

The theory states that phlogiston is responsible for combustion; phlogisticated substances are substances that contain phlogiston and are dephlogisticated when burned. Dephlogisticating is the process of releasing stored phlogiston, which is absorbed by the air during combustion. If air is absent, combustion cannot take place.

This theory has explanatory power. First, it explains why air is needed during combustion. Second, flames during combustion exhibit the very process of the escaping of phlogiston from the combustible materials into the air. Third, growing plants then absorb this phlogiston from air, which is why air does not spontaneously combust and also why plant matter (dried woods and trees) burns as well as it does.

This theory also has some predictive power. For example, substances that burned in air more easily were said to be richer in phlogiston; the fact that combustion soon ceased in an enclosed space was taken as clear-cut evidence that air had the capacity to absorb only a finite amount of phlogiston. When air had become completely phlogisticated it would no longer serve to support combustion of any material, nor would a metal heated in it yield a calx; nor could phlogisticated air support life through respiration. This is why in a sealed glass container, a fire or flame would eventually extinguish. Inhaling and exhaling by animals and humans was thought to take phlogiston in and out of the body, and this explains why the exhaled air does not support combustion.[22]

However, this theory has no “operational power” because phlogiston is not something alchemists can collect and use to create new combustible materials. Nonetheless, chemists did try to measure the weight of phlogiston by weighing combustible materials before and after combustion and found that phlogiston can have either a positive weight or negative weight, depending on the phlogisticated materials.

The phlogiston theory became widely accepted by the scientific community during the 17th and 18th centuries and also evolved over time to become more and more sophisticated and seemingly more “scientific.” In fact, this theory was so popular and so “successful” in explaining the phenomenon of combustion and gunpowder exploding, even the 18th century’s greatest scientists, such as Newton, Euler, Lagrange, Boyle, Hoke, and Laplace, could not refuse or refute the phlogiston theory.

---

[21] To gain a sense of the historical timing, remember that in 1665 Isaac Newton invented his calculus, in 1668 Isaac Newton constructed the first reflecting telescope, in 1673 Gottfried Leibnitz invented his calculus independently, in 1674 Anton van Leuwenhoek invented the compound microscope, in 1675 Isaac Newton invented an algorithm for the computation of functional roots, in 1678 Christian Huygens stated his principle of wave front sources, in 1684 Isaac Newton proved that planets moving under an inverse-square force law will obey Kepler's laws, in 1687 Isaac Newton published his *Principia Mathematica*, in 1690 Jacques Bernoulli solved the isochrone problem, in 1691 Jean Bernoulli solved the catenary problem, in 1697 Jean Bernoulli solved the brachistochrone problem, in 1698 Thomas Savery built a steam-powered water pump for pumping water out of mines, and in 1704 Isaac Newton published *Opticks.* This means that the development of chemistry was significantly lagging that of physics and mathematics.

[22] See https://en.wikipedia.org/wiki/Phlogiston_theory.

For example, the well-known chemist Joseph Black's student Daniel Rutherford discovered nitrogen in 1772 and the two chemists, without knowing what it was, used the phlogiston theory to explain such results. The residue of air left after burning or exploding gunpowder, in fact a mixture of nitrogen and carbon dioxide, was referred to by them as phlogisticated air, having taken up all of the phlogiston and being incapable of supporting fire.[23]

Conversely, when the famous British chemist Joseph Priestley discovered oxygen, he believed it to be dephlogisticated air, capable of combining with more phlogiston and thus supporting combustion for longer than ordinary air, as well as supporting respiration of animals. Carbon dioxide exhaled by animals was called "fixed air" or phlogisticated air because it does not support combustion or respiration.

Stahl considered phlogiston to be something like the light that cannot be put into a bottle but nonetheless has weight. In his phlogiston theory, burning of sulfur and other substances as well as calcination of metals involve loss of phlogiston by these substances. The metal had phlogiston, but the calx did not. But when the calx is reheated with carbon or other substances containing phlogiston, the calx regains phlogiston, restoring the original metal. Wood and charcoal were supposed to be rich in phlogiston.

During this time few chemists were doing chemical experiments using careful measurements, so phlogiston theory seemed adequate as a "scientific explanation" for the phenomenon of combustion. The theory "explained" a wide range of phenomenon and "correctly" described many scientific experiments. However, there were troublesome cases or contradicting evidences to the theory as well.

For example, on the "success" side, the Phlogiston theory seemed to give a good model for explaining many experiments:[24]

a. It accounted for the similar properties of metals, because they all contained phlogiston.

b. The metals and their calxes were related; they just had different amounts of phlogiston.

c. It accounted for why candles go out when placed in a closed jar. The air becomes saturated with phlogiston.

d. A mouse dies in a closed container, or in a container where a candle has been burned until it goes out, because the air is saturated with phlogiston.

e. Charcoal leaves very little ash when it burns because charcoal is nearly pure phlogiston by definition.

f. Some metal ashes convert back to metals when heated with charcoal because the charcoal restores the phlogiston to the ash.

But on the "failure" side, there were some troublesome problems and unaccountable cases. The most perplexing problem was this: Charcoal, when burned, loses almost all its weight, leaving only a light ash.

---

[23] See https://en.wikipedia.org/wiki/Phlogiston_theory.
[24] See https://www.acs.org/content/acs/en/education/whatischemistry/landmarks/josephpriestleyoxygen.html and https://www.lockhaven.edu/~dsimanek/ideas/consmass.htm.

So phlogiston was thought to have “positive weight.” But metals gained weight after combustion, as Robert Boyle had shown in the 17th century, so phlogiston was thought to have “negative weight.” So in one case, the charcoal loses weight when it loses phlogiston; but in the metal, it gains weight when losing phlogiston.

Does this positive and negative weight of phlogiston sound like the “positive” and “negative” TFP shocks proposed by Kydland and Prescott (1982)? This loss and gain in weight is indeed a “wedge” between the phlogiston theory and the data. But unlike neoclassical macroeconomists, this mysterious “wedge” was viewed by chemists as an uncomfortable factor within the phlogiston theory, while the Solow residual is viewed as a comfortable factor to macroeconomists and thus prompts them to introduce more and more shocks and wedges into a DSGE model (see, e.g., Smets and Wouters, 2007).

Some economists are tempted to interpret aggregate TFP shocks as manifestations of micro-level TFP shocks at the sectoral level or even at the firm level. But even if small firm-level TFP shocks can be magnified into aggregate movements, the same question remains as to why an industry or a firm can suddenly lose its technology or experience negative TFP shocks? Remember that even at the firm level economists still assume that the production function is given by $y = Af(k,n)$, where y is firm-level output, k is firm-level capital, n is firm-level employment, and A is firm-level TFP or the Solow residual.

The Chemical Revolution did not happen simply because of the negative weight of phlogiston. It happened because (i) chemists gradually discovered that the “air” is not a single element or pure substance, but there are many kinds of air with different properties: Some of them make combustion easier and some of them make combustion harder; (ii) all types of air or gas can actively participate in chemical reactions or be produced as products from the reactions; (iii) quantitative measurements were introduced into chemistry to measure the weight of inputs and outputs of a chemical reaction in closed containers, among other factors. [25]

The alchemist Jean Rey found in 1630 that tin gains weight by as much as 25% when it forms a calx. How could something gain weight if it loses phlogiston? Stahl, the founding father of phlogiston theory, did not believe negative weight so he rationalized this phenomenon of gaining weight cleverly, by suggesting that the weight increased because the phlogisticated air re-entered the metal to fill the vacuum left after the phlogiston escaped.

Similarly, macroeconomists rationalize negative technology shocks as slower technology growth, or sudden increases in oil prices in the Middle East, traffic jams on highway 90, bad weather at the North Pole, earthquakes in Southeast Asian, or sudden losses of human knowledge or firms’ organizational skills.

Some chemists entertained the notion that there might be two kinds of phlogiston, one with negative weight (called levity) and one with positive weight (called gravity). Influenced by the great success of Newton's work in mechanics, some chemists realized they needed to pay more attention to weight in their theories and experiments. So they extended the idea of positive weight and followed Isaac Newton and the Aristotle tradition by adding the idea of levity, or negative weight.

[25] See https://en.wikipedia.org/wiki/Chemical_revolution.

In the 18th century, most chemists were still doing chemical experiments in open containers, so they did not realize that the oxygen in the air was participating in the chemical reactions, until the Chemical Revolution detonated by the French chemist Antoine Lavoisier. That is, merely "weighing more carefully" the inputs and outputs in chemical experiments was not sufficient to resolve the difficulties facing the phlogiston theory, but conducting experiments in closed environments to include air as an active participant of chemical reactions was also critical. [26]

Phlogiston theory was eventually refuted or abandoned by chemists toward the end of the 18th century, thanks to the revolution lunched by Lavoisier. Several factors led to the Chemical Revolution.[27] First, there were the forms of gravimetric analysis that emerged from alchemy and new kinds of instruments that were developed in medical and industrial contexts. In these settings, chemists increasingly challenged hypotheses that had already been presented by the ancient Greeks. For example, chemists began to assert that there are more than four elements (fire, water, air, earth) in nature, in contrast to those Greek beliefs.

Second, earlier works by chemists such as Jan Baptist van Helmont helped to shift the belief that air existed as a single element to the one in which air is a mixture of distinct kinds of gasses. For example, the discovery of “fixed air” (carbon dioxide) by Joseph Black in the middle of the 18th century was particularly important because it empirically proved that “air” did not consist of only one substance and because it established “gas” as an important experimental substance. Nearer the end of the 18th century, the experiments by Henry Cavendish and Joseph Priestley further proved that air is not an element and is instead composed of several different gases.

Third, Lavoisier changed the field of chemistry by keeping meticulous balance sheets in his research, attempting to show that through the transformation of chemical species during any chemical reactions, such as combustion, the total amount of substance was conserved.

In 1789 Lavoisier published his *Traité Élémentaire de Chimie* (*Elementary Treatise on Chemistry*) in which he synthesized the work of others and his own in the past and coined the term "oxygen." Lavoisier's work was characterized by his systematic determination of weights and his strong emphasis on precision and accuracy. Lavoisier proposed to use oxygen as the explanation of the phenomena of combustion and respiration of animals. He also proved that water is composed of oxygen and hydrogen instead of a single element.

Lavoisier’s oxygen theory of combustion turned the phlogiston theory upside down. According to Lavoisier, the process of combustion is a process of oxidization in which the gas of oxygen contained in the surrounding air is combined with other chemical elements to form new forms of compounds. For example, Lavoisier explained that when phosphorus was burned, it combined with a large quantity of oxygen in the air to produce the acid spirit of phosphorus, and that the phosphorus increased in weight upon burning because it absorbed oxygen. Lavoisier applied the same principle to the burning of sulfur and conjectured in 1772 that "the increase in weight of metallic calces is due to the same cause."[28] Lavoisier also recognized that the “fixed air” discovered by Joseph Black was identical to the air emitted when metal calces were reduced with charcoal, or exhaled by animals and humans, and suggested that

[26] See https://en.wikipedia.org/wiki/Chemical_revolution.
[28] See https://en.wikipedia.org/wiki/Antoine_Lavoisier#Oxygen_theory_of_combustion.
[28] See https://en.wikipedia.org/wiki/Antoine_Lavoisier#Oxygen_theory_of_combustion.

the air which combined with metals on calcination and increased the weight might be Black's fixed air, that is, carbon dioxide (CO_2).

Lavoisier carried out systematic experiments on the calcination of tin and lead in sealed vessels and repeated the experiment of the English chemist Joseph Priestley by heating the red calx of mercury with a burning glass to generate a special type of air (oxygen). He conclusively confirmed that the increase in weight of metals in combustion was due to combination with oxygen in the air, which supported combustion. Lavoisier showed that the mercury calx was a true metallic calx in that it could be reduced with charcoal, giving off Black's fixed air (CO_2) in the process. When reduced without charcoal, it gave off oxygen, which supported respiration and combustion in an enhanced way. [29]

Lavoisier also conducted a series of experiments on the composition of water to prove that the so-called "oxygen" is measurable, re-creatable, and manipulable. Many chemists had been experimenting with the combination of the so-called "inflammable air" (first discovered by the British chemist Henry Cavendish), which Lavoisier termed hydrogen, with oxygen by electrically sparking mixtures of the gases. In fact, all of the researchers noted Cavendish's production of pure water by burning hydrogen in oxygen, but they interpreted the reaction within the framework of phlogiston theory. Lavoisier learned of Cavendish's experiment in June 1783 and immediately recognized water as the oxide of a hydroelectric gas.

Lavoisier synthesized water by burning jets of hydrogen and oxygen in a bell jar over mercury. The quantitative results were good enough to support the hypothesis that water was not an element, as had been thought for over 2,000 years by ancient Chinese and Greeks, but a compound of two gases, hydrogen and oxygen.

Lavoiser's *Elementary Treatise on Chemistry* laid out the core of his new chemistry—the oxygen theory, which became a most effective vehicle for the Chemical Revolution. It presented a unified view of new theories of chemistry, contained a clear statement of the law of conservation of mass, and denied the existence of phlogiston. This book clarified the concept of an element as a substance that could not be broken down by any known method of chemical analysis and presented Lavoisier's theory of the formation of chemical compounds from elements. It remains a classic in the history of science comparable to Newton's *Principia Mathematica* published one century earlier in 1687.

But, unlike Newton's classic mechanics, Lavoisier's chemistry was not based on mathematics and calculus, nor on the Euclidian axiom system of deductive reasoning, but rather on inductions and tireless experiments designed with the hope of discovering the precise causal chains of chemical reactions, the structure of compounds, and quantitative measurements of chemical reactions.

Despite the success of Lavoisier's oxygen theory of combustion, however, Lavoisier's anti-phlogistic approach remained unaccepted and even heavily criticized by many other chemists of his time. The fundamental reason was that scientists did not realize that phlogiston was not measurable, but that oxygen was measurable. Therefore, the notion of phlogiston is not refutable and quantifiable, while the notion of oxygen is refutable and quantifiable: The speed of combustion is faster the more oxygen is present. Once oxygen is used up or mixed with non-oxygen gases, combustion slows down. More

[29] See https://en.wikipedia.org/wiki/Chemical_revolution.

importantly, oxygen can be created or recreated from chemical reactions, while phlogiston cannot be captured or collected by any means.

4. What the Chemical Revolution tells us about the correct production function

As Paul Romer (2016) rightly criticized, the imaginary “shocks” or “wedges” have no micro-foundation themselves despite the “claim” by macroeconomists that the DSGE model is micro-founded. But this “claim” is also false. Even without the imaginary shocks, the neoclassical production function has no micro-foundation whatsoever and is precisely the reason for the need for imaginary shocks.

In other words, the most important reason for the prominent role of imaginary shocks and wedges in DSGE models is precisely the lack of a micro-foundation in the production function itself. Yet such production functions are used everywhere by both microeconomists and macroeconomists as well as econometricians.

So, not only the aggregate production function assumed by Robert Solow (1956) in his Nobel winning article is misspecified, but the disaggregate firm-level production function used by virtually all micro-economists and applied econometricians is also misspecified.

The reason is simple: Just like there are millions of different forms of chemical reaction equations, there are also millions of different forms of production processes. Although a theorist can identify the common patterns behind all types of chemical reactions so as to classify them into a few basic categories, there may also exist several basic types of production functions that characterize the millions of different production processes. But such a characterization or abstraction does not substitute for a concrete chemical reaction equation or a concrete production function in quantitative analysis.

For example, virtually all chemical reactions can be classified into five basic types: combination, decomposition, single-replacement, double-replacement, and combustion. Understanding these five types of reactions will be useful for predicting the general types of products of an unknown reaction, but such abstract characterizations do not substitute for specific chemical reaction equations in quantitative analysis, as shown below.

The first category of chemical reactions is the **combination reaction**, which is also known as a **synthesis reaction.** This is a reaction in which two or more substances combine to form a single new substance, such as two elements combining to form a compound. An example is solid sodium metal reacting with chlorine gas to produce solid sodium chloride. One common sort of combination reaction that occurs very frequently is an element reacting with oxygen to form an oxide. Metals and nonmetals both react readily with oxygen under most conditions.

The second category is the **decomposition reaction** in which a compound breaks down into two or more simpler substances. Most decomposition reactions require an input of energy in the form of heat, light, or electricity. The simplest kind of decomposition reaction is when a binary compound decomposes into its elements, such as when mercury oxide (a red solid) decomposes under heat to produce mercury and oxygen gas. A reaction is also considered to be a decomposition reaction even when one or more of the products is still a compound.

The third category is the **single-replacement reaction** in which one element replaces a similar element in a compound. For example, magnesium is a more reactive metal than copper, so when a strip of magnesium metal is placed in an aqueous solution of copper nitrate, it replaces the copper to form aqueous magnesium nitrate and solid copper metal. Many metals react easily with acids and one of the products of the reaction is hydrogen gas. For example, in a single-replacement reactio zinc reacts with hydrochloric acid to produce aqueous zinc chloride and hydrogen.

The fourth category is the **double-replacement reaction** in which the positive and negative ions of two ionic compounds exchange places to form two new compounds. Double-replacement reactions generally occur between substances in aqueous solution. In order for such a reaction to occur, one of the products is usually a solid precipitate, a gas, or a molecular compound such as water, such as when a few drops of lead nitrate are added to a solution of potassium iodide, a yellow precipitate of lead iodide immediately forms in a double-replacement reaction.

The fifth (last) category is the **combustion reaction** in which a substance reacts with oxygen gas, releasing energy in the form of light and heat. Combustion reactions must involve oxygen as one reactant. Many combustion reactions occur with a hydrocarbon, a compound made up solely of carbon and hydrogen. The products of the combustion of hydrocarbons are always carbon dioxide and water.

Notice that a chemical reaction can belong to more than one of the categories. Also, the five types of reactions are far from inclusive. There are in fact hundreds or even thousands of types of chemical reactions. For example, there are reactions called isomerization, neutralization, and hydrolysis. In an isomerization reaction, the structural arrangement of a compound is changed but its net atomic composition remains the same. In a neutralization reaction, an acid and a base react with each other and generally the product of this reaction is a salt and water. In a hydrolysis reaction, water is always involved.

The real-world production functions may also be classified into several abstract categories, such as the Cobb-Douglas production function or its CES (constant elasticity of substitution) generalization. However, such abstraction cannot substitute for any concrete presentations of different production processes, just like by classifying a gunpowder explosion as a combustion reaction or double-replacement reaction would be insufficient for understanding the actual reaction process of a gunpowder explosion. In other words, if a chemist simply calls the reaction of gunpowder a double-replacement reaction without providing the specific chemical reaction equation, such abstract knowledge would be highly insufficient for understanding the mechanisms behind a gunpowder explosion and quantitatively predicting its outcome, let alone helping create or discover new formulas of explosive materials.

Hence, it is precisely the neoclassical tradition of using highly abstract forms of production functions in all sorts of applied economic analysis in addition to highly misspecified utility functions and market structures that have led to the emergence of the ghost-like “imaginary shocks” and “wedges.”

In fact, the efforts to understand the precise input-output structure involved in a gunpowder explosion were the single most important factor motivating Lavoisier and many natural philosophers in his time to study combustion, which eventually led to the Chemical Revolution.[30]

Based on Lavoisier's new chemistry, a simple chemical equation for the combustion of black gunpowder is given by

$$2\ KNO_3 + S + 3\ C \rightarrow K_2S + N_2 + 3\ CO_2 \qquad (6)$$

This chemical reaction involves all five reaction types listed above. It suggests that gunpowder is a mix of three different components: potassium nitrate KNO_3, charcoal *C*, and sulphur *S.* Each of these components plays an important role in the combustion of gunpowder. Potassium nitrate, also known as saltpetre or saltpeter, decomposes at a high temperature to provide oxygen (O_2) for the reaction. This means that gunpowder doesn't need to be exposed to air to burn—and is why smothering fireworks will not stop them from burning. The charcoal is often represented simply as a source of carbon (*C*), which acts as a fuel, though it is actually a broken-down form of cellulose, with the approximate empirical formula C_7H_4O. Finally, the sulphur (S) can also act as a fuel, though its inclusion has more to do with the fact that it undergoes exothermic reactions (reactions that give off heat) at relatively low temperatures, providing more energy and lowering the ignition temperature of the charcoal.

A gunpowder explosion produces several products; none of them is the "phlogiston" imagined by chemists during the 17th to 18th centuries. The main products are potassium sulphide K_2CO_3, carbon dioxide CO_2, and nitrogen gas N_2 (in molecular form).

Equation (6) can be viewed as the "production function" of a gunpowder reaction: It not only tells how many atoms, molecules, or "formula units" of each kind are involved in a reaction, it also indicates the *amount* of each substance involved. Equation (6) says that 2 KNO_3 *formula units* react with 1 S *atom* and 3 C *atoms* to give 1 K_2S *formula unit*, 1 N_2 *molecule*, and 3 CO *molecules*. It also says that 2 *mol* KNO_3 would react with 1 *mol* S and 3 *mol* C, yielding 1 *mol* K_2S, 1 *mol* N_2, and 3 *mol* CO.

The balanced equation does more than this, though. It also tells us that $2 \times 2 = 4$ mol KNO_3 will react with $2 \times 3 = 6$ mol S, and that $\frac{1}{2} \times 2 = 1$ mol KNO_3 requires only $\frac{1}{2} \times 3 = 1.5$ mol C. In other words, the equation indicates that exactly 1 C must react *for every* 2 mol KNO_3 consumed. Therefore, for the purpose of calculating how much C is required to react with a certain amount of KNO_3, the significant information contained in Equation (6) is the *ratio*.

In fact, there is no simple equation for the combustion of gunpowder because the products, as well as the reactants, are numerous and varied depending on the purity of the raw materials as well as on the corning process and actual environment of the explosion. So, a more accurate, but still simplified, equation is given by

$$10\ KNO_3 + 3\ S + 8\ C \rightarrow 2\ K_2CO_3 + 3\ K_2SO_4 + 6\ CO_2 + 5\ N_2 \qquad (7)$$

[30] Lavoisier lunched the Chemical Revolution when serving as the director of French Government Bureau of Gunpowder and Saltpeter.

Once the exact chemical compounds involved in the gunpowder reaction and the oxidization principle are known, chemists can predict to a high level of accuracy the type of end products from the reaction, based on knowledge of the oxidization principle and atoms' chemical properties.

So, if we view the chemical reaction equation as a "production function" that relates chemical inputs to chemical outputs, then any misspecification of this "production function" will result in a "Solow residual" or "TFP wedge" between the actual output (products) of the chemical reaction and the model-implied output. For example, the phlogiston theory predicted that the output of a gunpowder explosion is only composed of ashes (charcoal deprived of phlogiston) and "phlogisticated air," anything else (such as the nitrogen gas and potassium sulphide) would be treated as unobservable "Solow residuals" or "TFP shocks."

5. What cellular respiration tells us about macroeconomic modeling

The importance of getting the neoclassical production function right in quantitative economic analysis can be further illuminated by the following example from biology, which illustrates how crucial the Lavoisier Chemical Revolution and the emphasis on quantifiable measurement of all inputs and outputs of a chemical reaction is to the development of biology. Understanding cellular respiration or metabolism in organic cells would be impossible had the phlogiston theory not been abandoned.

In a sense, the biological cell is like a firm or a national economy: Its cellular respiration process resembles the internal working structure of a firm or a national economy. Any misspecification of the chemical reaction processes in the cell's respiration mechanism would result in Solow-residual-like "wedges" that resemble our ignorance or the failure of the science.

But economics in the past century, especially in the past 30 years, has not achieved the same precision in describing the "metabolism" of firms or industries as biology has in describing cellular respiration.

Cellular respiration can occur either aerobically (using oxygen) or anaerobically (without oxygen). During *aerobic cellular respiration*, glucose reacts with oxygen, forming adenosine triphosphate (ATP) that can be used as a source of energy by the cell.[31] Carbon dioxide and water are created as byproducts. The highly simplified overall equation for aerobic cellular respiration is

$$C_6H_{12}O_6\ (s) + 6\ O_2\ (g) \rightarrow 6\ CO_2\ (g) + 6\ H_2O\ (l) + \text{ATP (energy)} \qquad (8)$$

where "s" denotes solid, "g" denotes gas, and "l" denotes liquid, as in any chemical equation. In cellular respiration, glucose and oxygen react to form ATP. Water and carbon dioxide are released as byproducts. The three stages of aerobic cellular respiration are *glycolysis* (an anaerobic process), *the Krebs cycle,* and *oxidative phosphorylation*.

Equation (8) can be viewed as an abstract "production function" of cellular respiration, but the devil is in the detail. More specifically, **cellular respiration** is a set of metabolic reactions and processes that take

[31] See https://en.wikipedia.org/wiki/Cellular_respiration.

place in the cells of organisms to convert chemical energy from oxygen molecules and nutrients (from food) into ATP and then release waste products (such as water and carbon dioxide). The reactions involved in respiration are catabolic reactions, which break large molecules into smaller ones, releasing energy because weak high-energy bonds, in particular in molecular oxygen, are replaced by stronger bonds in the waste products. Respiration is one of the key ways a cell releases chemical energy to fuel cellular activity. The overall reaction occurs in a series of biochemical steps, some of which are redox reactions. [32]

Although cellular respiration is technically a combustion reaction, it clearly does not resemble one when it occurs in a living cell, because of the slow, controlled release of energy from the series of reactions.

Nutrients that are commonly used by animal and plant cells in respiration include sugar, amino acids and fatty acids, and the most common oxidizing agent providing most of the chemical energy is molecular oxygen (O_2). The chemical energy released during respiration is stored in ATP, which can then be used to drive processes requiring energy, including biosynthesis, locomotion, or transport of molecules across cell membranes. Building on such "micro-foundations," the organs of animals and humans at a more aggregated level can thus function the way they do.

But notice that there does not exist a so-called "aggregate cell" with internal working mechanisms that exactly mimic cellular respiration, because animal cells and animal organs have very different structures and functions to perform, although the organ is composed of many cells. Only the Newtonian mechanical world view would view an animal organ simply as an aggregation or random collection of cells. Analogously, in economic theory, even if the firm-level Cobb-Douglas production function is a "correct" specification of the firm-level production function, the so-called aggregate production function invented by Robert Solow (1956) that looks identical to the firm-level production function does not really exist. Hence, any macroeconomic analysis based on such an imaginary aggregate production function is doomed to encounter troubles in quantitative analysis.

Consider **aerobic respiration first.** As in any combustion process, aerobic respiration requires oxygen (O_2) in order to create the energy-storage compound ATP. Although carbohydrates, fats, and proteins are consumed as reactants, aerobic respiration is the most efficient or "productive" method of breaking down the molecular pyruvate in glycolysis.

Most of the ATP produced by aerobic cellular respiration is made by oxidative phosphorylation. The energy contained in O_2 is released and used to create a chemiosmotic potential (a form of energy) by pumping protons across a membrane. This potential is then used to drive ATP synthase and produce ATP from adenosine diphosphate (ADP) and a phosphate group.[33] Biology textbooks often state that 38 ATP molecules can be made per oxidized glucose molecule during cellular respiration (2 from glycolysis, 2 from the Krebs cycle, and about 34 from the electron transport system).[34]

---

[32] See https://en.wikipedia.org/wiki/Cellular_respiration.

[33] ADP, also known as adenosine pyrophosphate (APP), is an important organic compound in metabolism and is essential to the flow of energy in living cells.

[34] However, this maximum yield is never quite reached, because of losses due to leaky membranes as well as the cost of moving pyruvate and ADP into the mitochondrial matrix; current estimates range around 29 to 30 ATP per glucose.

Aerobic metabolism is up to 15 times more efficient than anaerobic metabolism (which yields 2 molecules of ATP per 1 molecule of glucose) because the double bond in O_2 has more energy than other double bonds or pairs of single bonds in other common molecules in the biosphere. The two types of metabolism share the initial pathway of glycolysis, but aerobic metabolism continues with the Krebs cycle and oxidative phosphorylation. The post-glycolytic reactions take place in the mitochondria in eukaryotic cells and in the cytoplasm in prokaryotic cells.

The glycolysis process is a metabolic pathway that takes place in the cytosol of cells in all living organisms. Glycolysis can be literally translated as "sugar splitting" and occurs with or without the presence of oxygen. In aerobic conditions, the process converts one molecule of glucose into two molecules of pyruvate (pyruvic acid), generating energy in the form of two net molecules of ATP. Four molecules of ATP per glucose are actually produced; however, two are consumed as part of the preparatory phase. The initial phosphorylation of glucose is required to increase its reactivity (decrease its stability) in order for the molecule to be cleaved into two pyruvate molecules by the enzyme aldolase. During the pay-off phase of glycolysis, four phosphate groups are transferred to ADP by substrate-level phosphorylation to make four ATP, and two NADH are produced when the pyruvate is oxidized. The overall reaction can be expressed this way: [35]

$$\text{Glucose} + 2\ NAD^+ + 2\ P_i + 2\ ADP \rightarrow 2\ \text{pyruvate} + 2\ H^+ + 2\ NADH + 2\ ATP + 2\ H^+ + 2\ H_2O + \text{energy} \quad (9)$$

In the above chemical equation or "production function", starting with glucose, 1 ATP is used to donate a phosphate to glucose to produce glucose 6-phosphate. Glycogen can be converted into glucose 6-phosphate as well with the help of glycogen phosphorylase. During energy metabolism, glucose 6-phosphate becomes fructose 6-phosphate. An additional ATP is used to phosphorylate fructose 6-phosphate into fructose 1,6-bisphosphate with the help of phosphofructokinase. Fructose 1,6-biphosphate then splits into two phosphorylated molecules with three carbon chains that later degrade into pyruvate.

Pyruvate is oxidized to acetyl-CoA and CO_2 by the pyruvate dehydrogenase complex (PDC). The PDC contains multiple copies of three enzymes and is located in the mitochondria of eukaryotic cells and in the cytosol of prokaryotes. In the conversion of pyruvate to acetyl-CoA, one molecule of NADH and one molecule of CO_2 is formed.

During the metabolic process, there involves a process called citric acid cycle, which is also called the ***Krebs cycle*** or the *tricarboxylic acid cycle*. When oxygen is present, acetyl-CoA is produced from the pyruvate molecules created from glycolysis. Once acetyl-CoA is formed, aerobic or anaerobic respiration can occur. When oxygen is present, the mitochondria will undergo aerobic respiration, which leads to the Krebs cycle. However, if oxygen is not present, a different chemical process—fermentation of the pyruvate molecule—will occur.

[35] See https://en.wikipedia.org/wiki/Cellular_respiration.

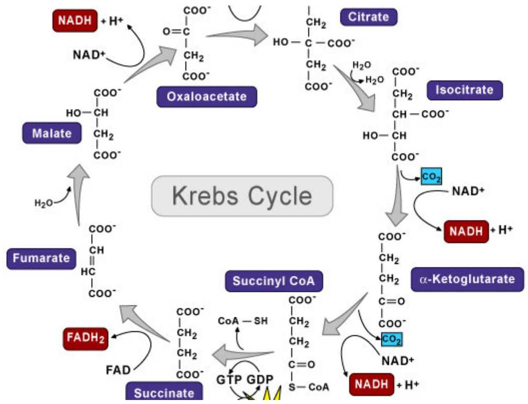

Figure 1. The Krebs Cycle.

In the presence of oxygen, when acetyl-CoA is produced, the molecule then enters the citric acid cycle (Krebs cycle) inside the mitochondrial matrix and is oxidized to CO_2 while at the same time NAD is reduced to NADH. NADH can be used by the electron transport chain to create more ATP as part of oxidative phosphorylation. To fully oxidize the equivalent of one glucose molecule, two acetyl-CoA must be metabolized by the Krebs cycle. Two low-energy waste products, H_2O and CO_2, are created during this cycle.

More specifically, the citric acid cycle is an 8-step process involving 18 different enzymes and co-enzymes. During the cycle, acetyl-CoA (2 carbons) + oxaloacetate (4 carbons) yields citrate (6 carbons), which is rearranged to a more reactive form called isocitrate (6 carbons). Isocitrate is modified to become α-ketoglutarate (5 carbons), succinyl-CoA, succinate, fumarate, malate, and, finally, oxaloacetate.

The net gain from one Kreps cycle is 3 NADH and 1 $FADH_2$ as hydrogen (proton plus electron)-carrying compounds and 1 high-energy Guanosine-5'-triphosphate (GTP),[36] which may subsequently be used to produce ATP. Thus, the total yield from 1 glucose molecule (2 pyruvate molecules) is 6 NADH, 2 $FADH_2$, and 2 ATP.

Although there is a theoretical yield of 38 ATP molecules per glucose during cellular respiration, such conditions are generally not realized, because of losses of energy such as the cost from moving pyruvate, phosphate, and ADP into the mitochondria. All are actively transported using carriers that utilize the stored energy in the proton electrochemical gradient.

Without oxygen, pyruvate (pyruvic acid) is not metabolized by cellular respiration but undergoes a process of fermentation. The pyruvate is not transported into the mitochondrion, but remains in the cytoplasm, where it is converted to waste products that may be removed from the cell. This serves the purpose of oxidizing the electron carriers so that they can perform glycolysis again and removing the excess pyruvate. Fermentation oxidizes NADH to NAD^+ so it can be reused in glycolysis. In the absence of

[36] GTP is a purine nucleoside triphosphate. It is one of the building blocks needed for the synthesis of RNA during the transcription process. Its structure is similar to that of the guanosine nucleoside, the only difference being that nucleotides like GTP have phosphates on their ribose sugar.

oxygen, fermentation prevents the buildup of NADH in the cytoplasm and provides NAD^+ for glycolysis. This waste product varies depending on the organism. In skeletal muscles, the waste product is lactic acid. This type of fermentation is called lactic acid fermentation. In strenuous exercise, when energy demands exceed energy supply, the respiratory chain cannot process all of the hydrogen atoms joined by NADH.

During anaerobic glycolysis, NAD^+ regenerates when pairs of hydrogen combine with pyruvate to form lactate. Lactate formation is catalyzed by lactate dehydrogenase in a reversible reaction. Lactate can also be used as an indirect precursor of liver glycogen. During recovery, when oxygen becomes available, NAD^+ attaches to hydrogen from lactate to form ATP. In yeast, the waste products are ethanol and carbon dioxide. This type of fermentation is known as alcoholic or ethanol fermentation. The ATP generated in this process is made by substrate-level phosphorylation, which does not require oxygen.

Fermentation is less efficient at using the energy from glucose: Only 2 ATP are produced per glucose, compared to the 38 ATP per glucose nominally produced by aerobic respiration. This is because most of the energy of aerobic respiration derives from O_2, which has a relatively weak, high-energy double bond. Glycolytic ATP, however, is created more quickly. For prokaryotes to continue a rapid growth rate when they are shifted from an aerobic environment to an anaerobic environment, they must increase the rate of the glycolytic reactions. For multicellular organisms, during short burst of strenuous activity, because of slower aerobic respiration, muscle cells use fermentation to supplement ATP production, so fermentation may be used by a cell even before oxygen levels are depleted, as is the case in sports that do not require athletes to pace themselves, such as sprinting.

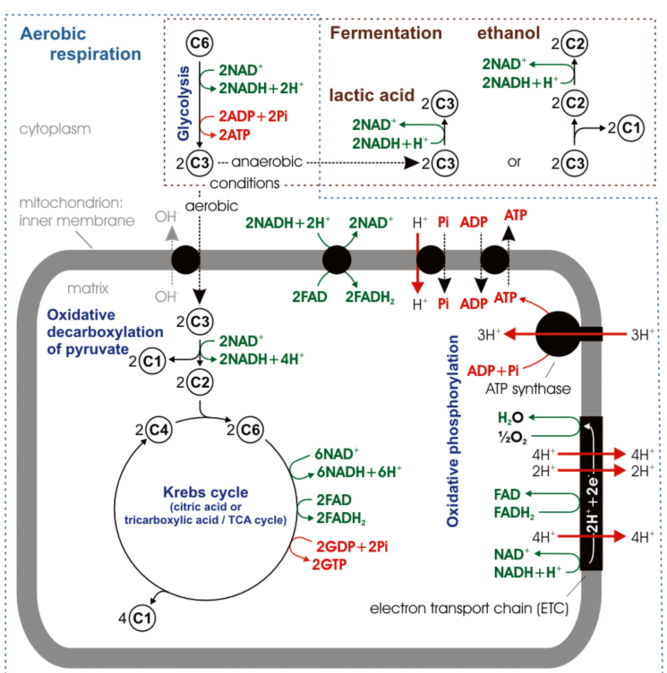

Figure2. Comparison of aerobic respiration and most known fermentation types in eukaryotic cells. Stoichiometry of aerobic respiration and most known fermentation types in eucaryotic cell. Numbers in circles indicate counts of carbon atoms in molecules, C6 is glucose $C_6H_{12}O_6$ and C1 is carbon dioxide CO_2. Mitochondrial outer membrane is omitted. (See https://en.wikipedia.org/wiki/Cellular_respiration.)

Now, imagine that a chemist were trained by economics to use a Cobb-Douglas production function to characterize both the aerobic respiration and fermentation by assigning two different TFP levels, *A > B*, to each production function:

$$Y = AF(K, N); \quad Y = BF(K, N); \tag{10}$$

where total output *Y = C + I* is composed of *C* (water) and *I* (ATP). Since ATP is also used during the production process, *K* can be interpreted as ATP and *I* can be interpreted as investment in *K*, and *N* denotes the nutrients such as sugar or glucose. The main difference between the two production functions in terms of the efficiency of production is that the technology coefficient *A* in aerobic respiration is greater than the technology coefficient *B* in fermentation.

Notice that these abstract production functions completely ignore the detailed causal mechanism of cellular respiration as well as the structures of both the cells and the chemical compounds used and produced in the entire metabolic process. These Arrow-Debreu-type production functions are complete black boxes where the Solow residuals {*A,B*} measure nothing but our ignorance. Any difference between the model-predicted outcomes and the actual outcomes would be captured and lump-summed into the imaginary Solow residuals.

Cellular respiration does have an abstract representation, as shown in equation (8), which is very much like our Cobb-Douglas production function but more structured. However, even this more structured abstract level of presentation is far from sufficient for scientists, as illustrated by Figure 1 and Figure 2. Yet in economics, we see that all quantitative models use the Cobb-Douglas production function, including applied empirical works in microeconomic analyses.

The consequence of relying only on an abstract presentation of cellular respiration in equation (8) is that if a poison or virus attacks the human body, there would be no way to understand the precise mechanisms behind or predict the consequences, let alone to prescribe medicine or treatment, which is exactly what happened during the Great Depression and the 2008 financial crisis: Although nobody predicted the crisis, every macroeconomist seemed to know how to “model” it ex post in an Arrow-Debreu framework. By introducing a large number of exogenous shocks to people’s tastes, firms’ Solow residuals, households’ borrowing constraints, or foreign demand, their DSGE models always explain the data 100%. But regarding how to avoid the next crisis, everybody shakes their head.

On the other hand, thanks to the Chemical Revolution that overthrew the phlogiston theory, chemists and biologists are in a much better position to deal with an analogous health crisis. For example, cyanide poisoning is an emergent crisis from exposure to cyanide. Symptoms include headache, dizziness, fast heart rate, shortness of breath, vomiting, seizures, slow heart rate, low blood pressure, and loss of consciousness; and the onset of symptoms usually occurs within a few minutes.

Because cyanide poisons the mitochondrial electron transport chain within cells and renders the body unable to derive energy (adenosine triphosphate-ATP) from oxygen, it can cause rapid death. Specifically, it binds to the a3 portion (complex IV) of cytochrome oxidase and prevents cells from using oxygen. This is why cyanide kills quickly; death occurs within seconds of a lethal dose of cyanide gas and within minutes of ingestion of a lethal dose of cyanide salt. The central nervous system and cardiovascular systems are chiefly affected.[37]

Cyanide poisoning is a form of histotoxic hypoxia because the cells of an organism are unable to create ATP; this is primarily due to the inhibition of the mitochondrial enzyme cytochrome c oxidase. Cyanide is quickly metabolized to 2-amino-2-thiazoline-4-carboxylic acid and thiocyanate, with a half-life of 10-30 minutes as a detoxifying mechanism.

Within a few hours of single ingestion, no cyanide can be detected, since all of it is metabolized, if death does not occur first. This thiocyanate has a long half-life (24 hours); is it typically eliminated through the kidneys, and its toxicity is about 0.01% (ten thousand times lower) than that of the cyanide parent molecule that it results from. [38]

Although medical science is still not perfect and there are still many diseases doctors still do not understand or know how to treat, it is in a much better position than economics in explaining phenomena, predicting counterfactuals, operating on patients, dealing with health crisis, and inventing new drugs.

6. Lessons for the micro-foundation of macroeconomics

Macroeconomics, on the other hand, has made little progress in becoming a quantitative science despite tremendous efforts in developing quantitative tools ever since the Keynesian revolution in 1930s and the rational-expectations revolution in the early 1980s. The unfortunate situation is that macroeconomics has become a branch of applied mathematics like classic mechanics at the time of Euler, Lagrange, and Laplace, instead of an empirical science like chemistry or biology. Becoming more and more mathematical and deriving everything from axioms instead of empirical observations is precisely where the problem lies. Chemistry and biology use little math, but this does not prevent them from being a powerful science. PhD students in economics programs study nothing but only mathematics and try to deduce economic laws from axioms, instead of drawing conclusions from the real world or human economic history by induction.

A common excuse we hear is that economists cannot conduct experiments, unlike chemists or biologists. This excuse is phony. There are ample observations and data that describe the actual working mechanism of a firm, an industry, and a nation. This information provides a clear picture of the inputs and outputs of a production process—exactly like the cellular respiration process—but is completely ignored by DSGE model builders.

For example, how steel is made in the real world clearly casts doubt on the accuracy of the Cobb-Douglas production function and its usefulness for quantitative analysis of economic fluctuations, but the actual production process of steel resembles very closely the chemical equation of combustion and the cellular respiration process of organisms. Yet macroeconomists have chosen to ignore such

[37] See, e.g., https://en.wikipedia.org/wiki/Cyanide_poisoning.
[38] See, e.g., https://en.wikipedia.org/wiki/Cyanide_poisoning.

information from the steel industry because they cannot apply fancy calculus to such a real-world production process.

Based on the article by Vicki Shipley (2017), [39] steel is an alloy primarily based on iron, but iron only occurs naturally as iron oxide in the earth's crust. Because of this, the ores must be "converted" or "reduced" using carbon. Carbon is produced after coal is coked and heated.

The coking process of coal involves heating metallurgical coal to 1000-1100 degrees Celsius in the absence of oxygen, creating hard, porous lumps called coke. The absence of air succeeds in driving out impurities, so the remaining carbon is pure. The entire process occurs over 12-36 hours in coke ovens.

The coke is either quenched immediately for storage or transferred directly into a blast furnace for iron-making. In the blast furnace, the coke is combined with iron ore and flux (small quantities of minerals like limestone), which collect more impurities. Heated air (about 1200 degrees Celsius) is blown to burn the coke and produce carbon monoxide, which reacts with the ore and melts the iron. The impurities are then drained, and voila, you have hot metal ready for steelmaking.

The most common steelmaking process occurs in a basic oxygen furnace and accounts for about 70% of the world's steel production. There, molten iron ore is combined with steel scraps and more flux. The scrap melts, the flux purifies, and the carbon content is reduced by 90%, resulting in liquid steel. Secondary processes may be applied afterwards, like the addition of elements such as boron or chromium.

Another 29% of the world's steel production is produced using electric arc furnaces. This process does not involve actual ironmaking, only the reuse of existing steel. These mini-mills run an electrical charge supplied by electrodes placed within the furnace, producing an arc of electricity that melts recycled steel scrap directly into new steel. Fluxes may be used to drive out impurities.

The additional 1% of steelmaking can be accounted for by less common methods like pulverized coal injection.

Molten iron from production furnaces typically runs through continuous casters and is formed into slabs, blooms, and billets. These basic products are further processed to create even more steel products through various operations such as hot rolling, cold rolling, and hot dip galvanizing.

A steel factory such as Pacesetter may process flat rolled steel from domestic and offshore mills to provide the customers with galvanized steel, cold rolled steel, galvannealed steel, aluminized steel, stainless steel, and galvanized bonderized steel products.

According to Shipley (2017), to survive in this highly competitive industry, a steel company often provides in-house services including high-speed precision slitting, precision blanking, and cut-to-length sheets, along with finishing touches such as pre-painting, embossing, perforating, and fabricating.

[39] See https://teampacesetter.com/everything-need-know-making-steel/

Typically, the products of a steel enterprise are used in a large number of industries, and their life cycle does not end there. Since steel is 100% recyclable, many of the products will someday return to the furnace, where they will be melted, produced, and finished all over again.

Steel production is the backbone industry of an industrial economy. Historically, the United Kingdom was the largest steel producer in the world by the second half of the 19th century and was then surpassed and replaced by the United States in the 20th century. Today, China has taken over as the industry leader. Why did this happen? The Solow residual in a DSGE model can tell us nothing.

Here is a typical DSGE presentation of the steel industry: $Y = AF(K, N)$, where Y denotes the quantity of steel produced and K and N denote capital and employment, respectively, needed in the steel industry. Obviously, this is not only a highly stylized but also a highly misspecified presentation of the true production process of steel, because it is too general to be informative and has been applied to represent any firm in any industry of any nation. Thus, the Solow residual is a critical black box to capture any gap due to such misspecification. Macroeconomists would argue that it is straightforward to make the production function more concrete, such as

$$Y_i = A_i F(x_1, x_2, , \cdots x_n;) \tag{11}$$

where the subscript *i* stands for industry *i* and x_i stands for the *i*th production factor used in producing output, which can be capital or labor or intermediate goods from other industries. Unfortunately, this does not solve the problem of missing information and misspecification. For example, imagine that one uses the following general representation of any chemical reaction equation:

$$\{X_i\} = \mathrm{AF}(\{X_j\}) \tag{12}$$

where *A* is a "phlogiston coefficient," {X} is the set of all chemical elements in the entire periodic table and all of their possible combinations (compounds) found in nature, $\{X_i\}$ and $\{X_j\}$ are subsets of {X}, and $F(\bullet)$ is a functional form indicating how the chemical elements or compounds are combined or separated during the reaction process. So the equation says that a general chemical reaction process is the transformation of a subset of chemical elements and compounds in the general set into a different subset of chemical elements and compounds in the general set. Any gap between the model-predicted outcome and the real outcome would be captured in the "phlogiston coefficient." This is surely very general, but too general to be informative and useful. Unfortunately, this is exactly what economists do in representing production functions across different firms, different industries, and different nations. They invented a production function that is too general to be informative and useful. Yet, they are also bold enough to assume a general elasticity of substitution between "capital" and "labor." In fact, such aggregate "capital" does not exist.[40]

Most importantly, macroeconomists would argue that the Solow residual is not a measure of our ignorance of the steel industry, but instead an exogenous shock that helps explain why the total steel output *Y* fluctuated over time and dramatically shrank in the United States in the past decades.

[40] See the Cambridge capital controversy at https://en.wikipedia.org/wiki/Cambridge_capital_controversy.

Clearly, such an abstract production function is not adequate for any quantitative analysis of the U.S. steel industry and the related business cycle during the Great Depression or the 2008 financial crisis, let alone a specific firm in this industry, yet the same production function is widely used for virtually all industries not only in quantitative macroeconomic analyses but also in applied microeconomic analysis in empirical works.

As Romer (2016) pointed out, a standard defense for the unrealistic assumptions in the Arrow-Debreu general-equilibrium model invokes Milton Friedman's (1953) methodological assertion that "the more significant the theory, the more unrealistic the assumptions."[41] In other words, economists are aware of the fact that "all of their models are false," but the statement "all models are false" "seems to have become the universal hand-wave for dismissing any fact that does not conform to the model that is the current favorite."[42]

There is no doubt that under certain circumstances and for certain types of questions, the Arrow-Debreu general equilibrium model is a very useful benchmark to gain **qualitative** insights. By the same reasoning, for certain types of questions, an abstract production function is enough and very helpful to demonstrate a point, such as to demonstrate the Ricardian principle of comparative advantage in international trade.

However, when the research question is to **quantitatively** explain and predict real-world growth or real-world economic fluctuations for a specific nation during a specific time frame, abstract models based on misspecified working hypotheses are no longer helpful, and they often misled researchers to arrive at the wrong answers. This is why economists still cannot adequately explain or predict real-world events such as the Great Depression, the 2008 financial crisis, and the sudden rise of China.[43]

Therefore, it is shocking to see that macroeconomists have been using highly misspecified or false models to conduct quantitative analyses and yet are comfortable with so many imaginary shocks and wedges floating around in their models.

But even for *qualitative economic analysis*, the working hypotheses underlying the Arrow-Debreu model must be constantly questioned, challenged, and tested, in order to advance the general theory and the field of economics and to see if they are suitable for the task of explaining real-world economic phenomena.

As discussed in the beginning in this paper, the two fundamental theorems of welfare economics are derived under a set of working hypotheses about human behaviors and social-economic structures that are hardly true in the real world, such as the following:

- Complete markets (i.e., markets for all imaginable goods and state-contingent assets exist and are complete; agents trading in these markets are infinitely lived with perfect rationality and without financial frictions such as borrowing constraints, or in the case of finite lives their altruistic parents are capable of taking good care of their offspring's welfare; and there exit perfect enforceability of contract so that a full set of financial tools can perfectly insure agents

[41] See Paul Romer (2016), p.5.
[42] See Romer (2016), p.5.
[43] See Wen (2015), "The Making of an Economic Superpower: Unlocking China's Secret of Rapid Industrialization."

against all types of idiosyncratic risks by making state-contingent plans against present and future uncertainties).

- Complete information (i.e., all market participants have perfect information on the market structures of the economy, including household preferences and firm production technologies and asset market trading rules, price signals, the quality of goods and services, the statistical distribution of exogenous shocks, and each other's actions and trading strategies).
- Price-taking behavior (i.e., all market participants behave "nicely" as price takers; their individual actions have no significant impact on prices; there is no cheating, collusion, robbery, stealing, price manipulation, monopoly power, or costs of entry and exit from market).
- No externalities (i.e., individuals' self-interested actions do not generate direct benefits or harm to other persons' productivity, happiness, or the ability to perform their market functions including information processing, and there do not exit public "goods" or "bads" that are essential or possibly destructive for production and market exchange, such as the need for infrastructure provision or violence prevention).
- No non-convexities in the utility functions, production technologies, or market structures (i.e., no increasing returns to scale for the division of labor or the specialization of consumption and production and no fixed costs for entering and existing markets, producing or consuming goods and services, organizing firms, creating production capacities, signing and enforcing contracts, or processing information or making decisions).

When any one or some of these highly idealized conditions are not met, a pure laissez-faire market economy not only does not achieve efficient allocation of resources, but can also lead to malfunction, stagnation, poverty traps, inequality, prolonged unemployment, speculative bubbles, financial crises, self-fulfilling boom-bust cycles, coordination failures, markets full of fake or "lemon" goods, business fraud, monopoly, oligarchy, and even self-destruction.[44]

First, the welfare economic theorems ignore the social-political environment in which markets can properly function. Neoclassical economists overlook three of the most important cornerstones (pillars) of the free market: (i) political stability, (ii) social trust, and (iii) infrastructure. In fact, infrastructure determines the time-spatial shape of the market and the volume, direction, and speed of the flow of goods and services. Yet, the three pillars require state power to build, nurture, and protect them and are missing in neoclassical economic theories.

[44] See "*Leviathan*" by Thomas Hobbes (1588–1679) about the "war of all against all." For arguments of inefficient or undesirable outcomes of market systems, see the classical book of Karl Polanyi, "*The Great Transformation: The Political and Economic Origins of Our Time*," and many of Joseph Stiglitz's works, including his classical analysis on imperfect information (available at http://scholar.google.com/scholar?q=stiglitz&hl=en&as_sdt=0&as_vis=1&oi=scholart&sa=X&ei=OFiEVePLBdOCyQSSoa3QDg&sqi=2&ved=0CBsQgQMwAA.) For the so-called dynamic-stochastic-general-equilibrium analyses of various types of market failures and the consequent economic collapse and chaos and boom-bust cycles, see, e.g., Azariadis, Kaas, and Wen (2015), "Self-fulfilling Credit Cycles"; Benhabib, Wang, and Wen (2014), "Sentiments and Aggregate Demand Fluctuations"; Coury and Wen (2009), "Global Indeterminacy in Locally Determinate Real Business Cycle Models"; Pintus and Wen (2013), "Leveraged Borrowing and Boom-Bust Cycles"; and Wu and Wen (2014), "Withstanding the Great Recession like China"; among many others.

This fundamental connection between political stability and orderly market activities based on social trust and infrastructure explains why after democracy was immaturely adopted or *imposed* on developing nations, such as Afghanistan, Egypt, Iraq, Libya, Pakistan, Tunisia, Ukraine and other parts of Eastern Europe, it failed to bring economic prosperity in ways the new institutional theorists and Western politicians would have hoped or predicted. Instead, democracy brought anarchy, chaos, and even endless civil wars to these poor nations. Markets would never emerge without political stability, social trust, and infrastructure. Yet a safe and unified national market is the absolute prerequisite of the division of labor and the existence of cooperatives, organized trade, and financial contracts.

Hence, the welfare economic theorems ignore (or assume away) the prohibitive social-economic costs of creating markets. Markets, especially the mass market, which is itself a "public good," are extremely costly to create even under long-term political stability:

" [S]o it is upon the sea-coast, and along the banks of navigable rivers, that industry of every kind naturally begins to subdivide and improve itself, and it is frequently not till a long time after that those improvements extend themselves to the inland parts of the country…. There could be little or no commerce of any kind between the distant parts of the world…. Since such, therefore, are the advantages of water-carriage, it is natural that the first improvements of art and industry should be made where this convenience opens the whole world for a market to the produce of every sort of labor, and that they should always be much later in extending themselves into the inland parts of the country. The inland parts of the country can for a long time have no other market for the greater part of their goods, but the country which lies round about them, and separates them from the sea-coast, and the great navigable rivers. The extent of the market, therefore, must for a long time be in proportion to the riches and populousness of that country, and consequently their improvement must always be posterior to the improvement of that country." (Adam Smith, *The Wealth of Nations*, 1776, Chapter III)[45]

Geographical isolation and distances are not the only obstacles to the formation of markets. Think of social distances and social isolation. For example, the 2.5 million people of Papua New Guinea in the early 1970s had about 700 regional indigenous languages. Some, like Abaga, were spoken by as few as five people.

It is perhaps equally if not more costly to create market regulatory institutions to prevent cheating and fraud. Human beings do not always behave "nicely." The quickest way to meet consumption needs or generate wealth is not through hard labor; it can be much simpler to steal or seize other people's goods and wealth. So market participants may cheat, collude, lie, and/or steal. The rule of law may apply to these actions, but the rule of law means little when it is not enforceable. Enforcement is itself extremely costly and often a fundamental breeding ground for corruption. Human beings possess both creative and destructive powers that they can impose on others, and in the worst cases the results can be fatally harmful. It can cost a person very little effort to save a life (business) or destroy a life (business). Market

[45] Adam Smith also mentioned about the other types of costs of conducting market exchanges, such as the costs of preventing robbery and piracy. Indeed, one of the most important functions of the powerful British navy was to protect its maritime trade. Mass international trade was impossible without a strong state and its military projection capacity (even true in today's "peaceful" world).

forces need not be exclusively creative or destructive. They can be creative only under proper regulations but destructive without such regulations.[46]

Fourth, human beings are endowed with finite physical and intellectual abilities. They have only two legs to paddle a waterwheel and two arms to spin yarn and only a limited number of brain cells to learn and process information. So it is in their best interest to cooperate to accomplish tasks, conduct business, and be competitive in the market (based on the Smithian principle of the division of labor). Yet, cooperation is costly as well (extremely so in the early stage of development in agrarian societies), and market principles fail to apply to collective activities, such as activities within cooperatives, which are ruled by the usual hierarchical power structures that have existed even in ancient civilizations before any modern markets and organizations appeared. In fact, the Europeans in the 16th to 19th centuries were so much better at military organization than the American Indians or African slaves because the Europeans were constantly at war against each other at least since the First Crusade.

Also, for much of history, price mechanisms and spot-market bargaining as well as democracy have not existed within any enterprises—much as they do not exist in the military today. No teamwork within any firm can be bargained on site with price tags. No CEOs or generals are democratically elected by employees or soldiers. What this suggests is that the so-called "market mechanism" described by the Arrow-Debreu model is not a universal mechanism for resource allocation.

Last but not least, physical and intellectual abilities do not distribute equally among human beings. Hence, a free market where the "winner takes all" can lead to dismal outcomes and excessive poverty instead of prosperity.

In light of these problems, the working hypotheses of the Arrow-Debreu general-equilibrium paradigm are troublesome. Yet the fundamental belief on the efficiency of a market economy is built on such a research paradigm. In reality, the market is only part of the resource-allocation mechanism even in modern developed nations and always complemented by non-market forces. Therefore, economic organizations, cooperatives, commune spirit, teamwork, ethics, trust, ideology, spirit, religion, culture, the state, and all types of coordinated and cooperative and collective actions are essential (in addition to markets) for achieving "efficient/effective" resource allocation and economic development.

What markets provides, among other things, is a form of impersonal competition and creative destruction, a discipline on management and technology adoption, and a mechanism of Darwinian "natural selection" of the "fittest." However, the fundamental limitations of human rationality,

---

[46] As a matter of fact, all states regulate, such as the United States. "From rules and laws governing trade, banking, and education to hazardous material, health standards and so on, the state rules on what will be produced, how it will be produced and often who will be the beneficiary of what… For example, in the United States of America the costs of social regulation tripled from $80 billion in 1997 to $267 billion in 2000….of the total $542 billion in regulatory costs in the United States of America (9 per cent of GDP in 1991 dollars), $189 billion were the costs associated with the paperwork and implementation of regulations." (Seema Hafeez, 2003, p. 1-3) Economic historian Marc Law and Sukkoo Kim (2011, p. 113) also wrote: "Despite the United States being the world's largest free market economy, government regulation of economic activity is a pervasive feature of the American economy…. The foods Americans eat, the cars they drive, the medicines they take, and the financial institutions from which they borrow and to which they lend are all subject to some kind of regulation." But ironically, regulations do not appear in neoclassical growth models and institutional theories, or in the case they do, they appear as negative constraints and impediment to development and growth.

information-processing capacity, foresight, and intellectual abilities dictate that the winners of market competition are not autarkic and anarchic individuals or artisan workshops, but rather well-organized corporations and nation-states based on *non-market* principles such as the division of labor, specialization, collaboration, command, commitment, dignity, ethics, friendship, honor, ideology, loyalty, reputation, shame, and trust.

Hence, markets and organization go hand in hand; the invisible hand and the visible hand go hand in hand; self-interests and collective interests go hand in hand; private property rights and public property rights go hand in hand; deregulation and regulation go hand in hand; freedom and discipline go hand in hand; and individuals and the state go hand in hand.

By denouncing or undermining the pivotal role of the government, the state, and the evolutionary institutions built upon the social collective spirit in economic activities, in market creation, in industrial organization, and in production and trade network formation, as well as the necessary market regulations, the Arrow-Debreu style laissez-faire approach to economic analysis is doomed to be falsified, along with its highly misspecified production function and the implied "imaginary wedges" in DSGE models.

Conclusion

The founding fathers of the Arrow-Debreu paradigm and its DSGE reincarnation were a group of mathematical geniuses well trained in physics and statistics. However, the activities of humans are more closely associated with animals than with particles. Hence, using the Newtonian mechanics as a role model for building economic theories can be misleading despite the heavy use of mathematics. The fields of chemistry and biology rarely rely on high-powered mathematics, yet they have traveled far in understanding the laws behind combustion and cellular respiration in organisms. Consequently, these branches of science are far more powerful in explaining, predicting, and changing the world than neoclassical economics.

The embarrassing inability of quantitative macroeconomics to explain the Great Depression, to predict the 2008 financial crisis, or to engineer economic prosperity or industrial revolution in developing countries is not because economists cannot do experiments with the economy. In fact, the long history of civilization and the Industrial Revolution has provided rich knowledge and repeated "experiments" and the detailed information about the actual production processes of firms and industries available to entrepreneurs and businessmen has been disregarded and ignored by economists as useful knowledge to build their models of growth and business cycles. One of the reasons seems to be associated with the founding fathers' preference for doing economic analysis like theoretical physicists, by siting in their offices with pencil and paper, rather than like chemists or biologists, by working in their labs. The lack of labs in economic research is not a good excuse, however, because firms, companies, corporations, industries, cities, counties, states, and nations are as good as chemical labs in providing information about the production function, the input-output relationship, balance sheets, management, the decision-making process, the accumulation of wealth, and trading networks. Adam Smith and John Maynard Keynes did not need "labs" to discover the division of labor and the circulation of money, for example.

Hence, the poverty of quantitative macroeconomics is the consequence of choice and ideology. Its elegant mathematical equations are normative rather than positive statements about commerce, production, capital, and wealth. If quantitative macroeconomics is to become a science, it needs Lavoisier, not Newton.

References

Azariadis, C., Kaas, L., & Wen, Y. (2016). Self-fulfilling credit cycles. *The Review of Economic Studies*, *83*(4), 1364-1405.

Acemoglu, D., & Robinson, J. A. (2012). *Why nations fail: The origins of power, prosperity, and poverty*. .

Benhabib, J., Wang, P., & Wen, Y. (2015). Sentiments and aggregate demand fluctuations. *Econometrica*, *83*(2), 549-585.

Blanchard, O. (2016). Do DSGE Models Have a Future? Peterson Institute of International Economics, PB 16-11.

Chari, V. V., Kehoe, P. J., & McGrattan, E. R. (2002). Accounting for the Great Depression. *American Economic Review*, *92*(2), 22-27.

Chari, V. V., Kehoe, P. J., & McGrattan, E. R. (2007). Business cycle accounting. *Econometrica*, *75*(3), 781-836.

Cole, H. L., & Ohanian, L. E. (1999). The Great Depression in the United States from a neoclassical perspective. Federal Reserve Bank of Minneapolis *Quarterly Review*, *23*, 2-24.

Cole, H. L., & Ohanian, L. E. (2004). New Deal policies and the persistence of the Great Depression: A general equilibrium analysis. *Journal of Political Economy*, *112*(4), 779-816.

Coury, T., & Wen, Y. (2009). Global indeterminacy in locally determinate real business cycle models. *International Journal of Economic Theory*, *5*(1), 49-60.

Hobbes, T. (1980). Leviathan (1651). *Glasgow 1974*.

Krugman, Paul (2009). How did economists get it so wrong? *New York Times Magazine*, September 2,

Kydland, F., & Prescott, E. (1982). Time to build and aggregate fluctuations. *Econometrica*, 50, 1345-1370.

Lin, J. Y. (2012). *New structural economics: A framework for rethinking development and policy*. The World Bank.

Mas-Colell, Andreu, Michael Dennis Whinston, and Jerry R. Green. *Microeconomic Theory*. New York: Oxford University Press, 1995.

North, Douglass C. (1990). *Institutions, Institutional Change and Economic Performance*. Cambridge: Cambridge University Press, 1990.

Pintus, P. A., & Wen, Y. (2013). Leveraged borrowing and boom-bust cycles. *Review of Economic Dynamics*, *16*(4), 617-633.

Polanyi, K. (2001). *The great transformation: The political and economic origins of our time*; foreword by Joseph E. Stiglitz. Beacon Press. Boston.

Romer, P. (2016). The trouble with macroeconomics. *The American Economist*, *20*, 1-20.

Ruan, J., & Zhang, X. (2009). Finance and cluster-based industrial development in China. *Economic Development and Cultural Change*, *58*(1), 143-164.

Shipley, Vicki, 2017, Everything You Need to Know About Making Steel, (https://teampacesetter.com/everything-need-know-making-steel/)

Smets, F., & Wouters, R (2007). Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach. *American Economic Review*, 93, 586-606.

Smith, A. (1776) "An inquiry into the nature and causes of the wealth of nations: Volume One." London: printed for W. Strahan; and T. Cadell, 1776.

Solow, R. M. (1956). A contribution to the theory of economic growth. *The Quarterly Journal of Economics*, *70*(1), 65-94.

Stiglitz, Joseph (2011), “Rethinking Macroeconomics: What Failed and How to Repair It” *Journal of the European Economic Association*, 9(4), pp. 591-645.

Stiglitz, J. E. (2018). Where modern macroeconomics went wrong. *Oxford Review of Economic Policy*, *34*(1-2), 70-106.

许成钢, (2020), “经济学理论的贫困”，《经济学（季刊）》2020 年 10 月第 20 卷第一期.

Wen, Y. (2016). *The Making Of An Economic Superpower: Unlocking China's Secret Of Rapid Industrialization*. World Scientific.

Wen, Y., & Fortier, G. E. (2019). The visible hand: The role of government in China’s long-awaited industrial revolution. *Journal of Chinese Economic and Business Studies*, *17*(1), 9-45.

Wen, Y., & Wu, J. (2019). Withstanding the Great Recession like China. *The Manchester School*, *87*(2), 138-182.