

NEW PERSPECTIVES ON POLITICAL ECONOMY

A bilingual interdisciplinary journal

Vol. 10, No. 1, 2014

CEVROINSTITUT
ACADEMIC **PRESS**



CEVROINSTITUT
[vysoká škola]

NEW PERSPECTIVES ON POLITICAL ECONOMY

A bilingual interdisciplinary journal / Vol. 10, No. 1, 2014

New Perspectives on Political Economy is a peer-reviewed semi-annual bilingual interdisciplinary journal, published since 2005 in Prague. The journal aims at contributing to scholarship at the intersection of political science, political philosophy, political economy and law. The main objective of the journal is to enhance our understanding of private property, market- and individual liberty-based perspectives in the respected sciences. We also believe that only via exchange among social scientists from different fields and cross-disciplinary research can we critically analyze and fully understand forces that drive policy-making and be able to spell out policy implications and consequences. The journal welcomes submissions of unpublished research papers, book reviews, and educational notes.

Published by CEVRO Institute Academic Press

EDITORIAL ADDRESS:

New Perspectives on Political Economy

CEVRO Institute, Jungmannova 17, 110 00 Praha 1, Czech Republic

Manuscripts should be submitted electronically. All manuscripts and correspondence should be addressed to nppe@vsci.cz.

Full text available via DOAJ Directory of Open Access Journals and also via EBSCO Publishing databases.

INFORMATION FOR AUTHORS

Authors submitting manuscripts should include abstracts of not more than 250 words and JEL classification codes. New Perspectives on Political Economy edits for clarity, brevity, and in accordance with the Chicago Manual of Style. Authors should use footnotes rather than endnotes or in-text references, and must include complete bibliographical information. Authors should include information on their titles and professional affiliations, along with e-mail address.

ISSN 1801-0938 (on-line)

ISSN 1804-6290 (print)

EDITOR-IN-CHIEF:

Josef Šíma, President, CEVRO Institute, Prague

MANAGING EDITOR:

Pavol Minárik, CEVRO Institute, Prague

ADVISORY EDITOR:

Ladislav Mrklas, CEVRO Institute, Prague

ASSOCIATE EDITORS:

Pavel Pšeja (political science), Masaryk University, Brno

Michal Kořán (international relations), Institute of International Relations, Prague

Alena Zemplerová (political economy), Institute of Economics, Academy of Science, Prague

Roman Cardal (political philosophy), CEVRO Institute, Prague

Jakub Kříž (law), CEVRO Institute, Prague

BOOK REVIEW EDITORS:

Jan Jireš CEVRO Institute, Prague

LANGUAGE EDITOR:

Shelly Gussis

EDITORIAL BOARD:

Terry Anderson, Hoover Institution,
Stanford University

Peter Boettke, George Mason University

Hardy Bouillon, Trier University

Enrico Colombatto, University of Turin

Christie Davies, University of Reading

Frank van Dun, Ghent University, Faculty
of Law

Richard Ebeling, Northwood University

Richard Epstein, New York University
School of Law

Robert Higgs, The Independent Review

Jesus Huerta de Soto, Universidad Rey

Juan Carlos

Jorg Guido, Hulsman Universität Anger

Zdeněk Kühn, Charles University, Faculty
of Law

Miroslav Novák, CEVRO Institute

Svetozar Pejovich, Texas A&M University

David Schmitz, University of Arizona

Gerald Steele, Lancaster University

LESSONS FROM DEVELOPMENTAL STRATEGIES OVER THE LAST HUNDRED YEARS¹

JAN WINIECKI²

This is a brief story of the most popular developmental strategies applied over the last century (to be more precise, since 1917), of their economic outcomes, as well as of the “conspiring circumstances” that led to adjustment or complete abandonment of these strategies. I chose two such popular strategies and, then, the third, much closer to classical economic prescriptions. In other words, a return to economic foundations.

What was the nature of these two popular strategies? Chronologically:

- In the (communist) East it was a so-called “steep ascent” strategy envisaging the fastest possible increase in the share of industry, and more specifically heavy industry. The choice of strategy took place under the circumstances of a complete economic system change. A centrally planned and administered economy was imposed in place of the – already highly constrained – market economy (the former system in the English language literature had more often been called a command economy).
- In the much more varied South (called after World War II the “Third World”) it was also a highly interventionist, *dirigiste* (to use a French term) developmental strategy of managing the economic development from the center, while accepting, however grudgingly, the existing, but strongly constrained, market economy as a system.

Both strategies aimed to catch up with the West’s level of economic development, but – thanks to the dominance of the state over the operations of the economy and hopefully over the future economic development – to do it in a shorter time span than did the West. To put it differently, they were both “shortcut” strategies. The idea was that saving time means also saving the cost of moving the economy from underdevelopment to development and reaching the West’s level of GNP *per capita* (nowadays a standard measure of development) faster than under the largely spontaneous, market order-determined development path.

1 This is a written version of a presentation offered at the 2014 Conference on Political Economy at CEVRO Institute in Prague.

2 Professor and Chair of International Trade and Finance at University of Information Technology and Management, Rzeszow, Poland. This term he is also a member of the Monetary Policy Council at the Polish National Bank.

The results were in the case of both strategies drastically different from expectations. With the benefit of hindsight, we now know that the goal of catching up was not achieved and strategies themselves were abandoned one way or another:

- Either in a complete change after communism's collapse in the 1989–91 period; or Over a longer period, and often through a series of partial reforms departing step-by-step from development economics' theorizing-based strategy.

Although the spirit of *dirigisme*, of tinkering with the economy is alive and well (even kicking recently, since the Great Financial Crisis), almost no one seriously suggests nowadays the return to central planning and administration, combined with the strategy of the rapid build-up of heavy industry (and sacrificing everything else, including the improvement in living standard of the population). No one suggests either the return of the “Permit raj”, a symbol of the almighty bureaucracy and planning for development in “India of the slow economic growth era” between 1947 and 1990. That is, India under the spell of the *dirigiste* development economic ideas, the believers of which pointed often to India as a shining example in the early post World War II period.

The idea that a “shortcut” strategy does not work has been revealed step-by-step, decade after decade. The same happened with the dawning realization that classical economic prescriptions did not lose their validity also with respect to poor countries. This realization was strengthened by the emergence of the new strands of economic thinking, underpinning classical economics, such as new economic history, property rights theory, public choice and similar strands of theorizing. The post-WW II revival of Austrian economics, crowned in 1974 with the Nobel Prize in economics for Friedrich Hayek, was also a part of the process in question.

CONSPIRING CIRCUMSTANCES

John Stuart Mill in the middle of the 19th century wrote that new ideas need “conspiring circumstances” to succeed. Mill's concept fits the post-World War II story of new 20th century orthodoxies, theorizing on economic development. The clash of theorizing and realities produced – over a period of decades – solid, and accumulating, evidence of multiple failures of these strategies. The evidence in question and its political *cum* economic consequences created a fertile ground for Millsian conspiring circumstances that led to the successful reemergence of classical market-based economic ideas.

Significantly, the accumulating evidence of interventionist failures became increasingly visible in more or less the same period all over the world: in the mid-1960s.

- In the most strictly controlled world of communism a series of reforms in Eastern Europe (as the Soviet Union and its dependencies were called at the time) failed to produce positive changes in the level of efficiency. Lasting improvements were neither visible in reforms in Hungary and Poland in the second part of 1950s, nor later in the USSR, Czechoslovakia, and communist East Germany. Costly distortions of

aberrantly high resource intensity, resistance to innovation and resultant increasingly obsolete product range, dramatically low quality of output, etc., continued unabated.

- Moreover, Czechoslovak reforms spilled over, under the people's pressure, into politics and resulted in the Soviet-led military invasion and the removal of the reformist government. The message was clear for Eastern Europe: no serious reforms were to be pursued, only tinkering. An ossification process of the already extremely rigid economic system began.
- With respect to less developed countries of the South, asaying became popular at the time: "They were first facing the problems of economic growth and now they face the growth of economic problems". And indeed they were. Costly distortions began to spread and became more painful with the shift from the "easy" import substitution phase to that of substituting imports with domestic production of heavy industries, where countries of the South did not possess comparative advantages.

Unsurprisingly, these distortions were roughly similar in both communist countries and the less developed South. After all, both "East" and "South" pursued the same type of a "shortcut" strategy: forced economic development led by the strong hand of the government (with a greater or smaller dose of central planning thrown in). With similar strategies, distortions were of the same kind. Where they were not, it was due to the fact that distortions in communist economies resulted not only from a "shortcut" strategy, but also from the economic system, prone to generate heavy distortions on its own, regardless of the strategy.

The West had faced, of course, a largely different range of economic problems, those of developed market economies. However, the post-WW II expansion of the role of the state, the interventionist macroeconomic governance and overregulation, also began to show with the passage of time increasingly adverse consequences. "One way" Keynesian demand management, that is expansionist fiscal policies at growth decelerations and almost no restrictive fiscal policies at excessive growth accelerations, created a pattern of ever higher inflation from one business cycle to another, beginning in the late 1950s. Simultaneously, economic growth tended to decrease from cycle to cycle.

And in 1970s, with the explosion in commodity prices, Western economies revealed another weakness, this time created by overregulation. The rigidity was so serious that a new term was coined: *eurosclerosis*. Let me add that although the term applied to Western Europe, the US was not that far behind.

To sum up, the foregoing developments from World War II to the mid-1960s and beyond to 1970s, when a variety of difficulties were revealed nearly everywhere, should have brought into attention the J.S. Mill's *dictum* that ideas need conspiring circumstances to succeed. The dictum applies to both theoretical domination and policy practice. It signaled that new orthodoxies did not work as expected and laid factual foundations for the future return of classical economics, strengthened by new, or facing the renewal similar strands of free market-based thinking.

THE LESSON FROM LITTLE DRAGONS

The return of the capitalist market with the corollary of free prices and private economic agents operating in a competitive environment had taken place step by step since the second half of the 1960s. The first step in that direction took place even before the start of the intellectual expansion of free market ideas in the post-WW II period. That step was made by a small group of East Asian countries called alternatively “little dragons” or “little tigers” to differentiate them from the big dragon, Japan.

Contrary to common perception, their choices in favor of the capitalist market were taken in different ways and in different environments. Hong Kong, under the energetic classical liberal economic administrator, pursued its famous policy of “active non-intervention” already since the mid-1950s. Taiwan and South Korea were typical *dirigiste* economies after their emergence as independent states, with all the failings of such a developmental strategy. They changed in late 1950s and early 1960s by allowing a much greater role of markets and greater openness *vis-à-vis* the world market. The change took place under pressure from the US, its defender in the military realm and the supplier of a very substantial amount of foreign aid. And Singapore threw its lot with more markets and more openness in mid-1960s, when it decided to become an independent city-state.

Although their paths toward more markets and more openness were different, the economic outcomes of the institutional changes were the same. Economic growth rates increased sharply, while manufacturing export growth rates tripled or quadrupled in comparison with the *dirigiste* past.

Interestingly, they for a long time were not seen as examples for other less developed economies of the South. The strength of beliefs in development economics-based ideas of managing from the center, planning investment and controlling foreign trade had been very strong in that period. Ideological preferences were, moreover, reinforced by interests of the ruling elites, which enjoyed both power and the ability to benefit materially, stemming from the management of the national economy from the center.

Thus, “little dragons” were first called “special cases”, with their critics stressing the alleged short-term adverse effects of “dangerous” market reforms. They needed

15 or more years of high growth and even higher exports to find a few – mostly timid – followers in the South.

Although a strong version of interventionism, associated with managing the economy from the center, began to lose gradually some of its attraction, the slowness of the process was not a natural process, but a process obstructed by an innovative rearguard action of intellectual supporters of *dirigisme*.

They first ignored economic successes of the little dragons, then called them special cases (not applicable to the wide and diverse group of Southern countries) and finally “adopted” the dragons as fellow interventionists. They announced triumphantly that “little dragons” also intervened, the only difference being that they intervened “wisely”.

However, the major difference was not that they intervened wisely (adoptionists never explained the secret of wise intervention), but that they intervened less. Besides, it is worth

keeping in mind that the world was a different place from the 1950s to the 1970s. Nobody, including the West, was then a paragon of excellence as far as freedom of markets was concerned. In fact, “little dragons” succeeded primarily by sticking to market fundamentals, like – for example – following the Heckscher-Ohlin version of classical trade prescriptions in their export choices.

The period from the mid-1960s to the late 1980s was a slow uphill march of a greater role of markets and greater openness in the South, a march with detours and returns at various times and in various places. Moreover, the intellectual discontent and the discontent with economic outcomes were less directly felt in the realm of politics, given the undemocratic governance regimes in a majority of countries in question.

Nonetheless, the number of those governments that both were *bona fide* reformers and felt brave enough to face the massive criticism from the still large numbers of believers in the tenets of development economics increased over time. Rather small in the 1970s, it grew in the 1980s, and multiplied in the later period.

THE WEST AND THE POST-COMMUNIST TRANSITION

The story of the West, although beset by different problems from those typical for economic development, was more straightforward politically. Being the only democratic area at the time helped. Intellectual discontent and economic outcome discontent converged in the second half of the 1970s and resulted in political changes through the ballot box.

The rest is history too well known to be repeated here. It is enough to state that Thatcher and Reagan free market “counterrevolutions” not only liberalized the respective economies. They also had an impact on the West as a whole. The 1980s and 1990s were for the Western world the age of deregulation and macroeconomic moderation (and the age of – less successful – attempts to put public expenditures in order). The United States and the United Kingdom benefited most in terms of economic growth. The former grew over the 1980–2005 period faster by almost one percentage point annually than the major West European economies, while the latter grew faster than Germany, France, and Italy. And it is worth keeping in mind that in the first 25 years after World War II, the growth pattern was just the opposite.

The communist system turned out to be the most resistant to both new ideas (suppressed in the public sphere) and discontent with economic outcomes (lied about but palpably visible in long queues and other manifestations of excess demand). The tight totalitarian control over the subjects was the cause. But once the rot started to spread, it spread more quickly due to the lack of *bona fide* defenders of the regimes. As economic historian David Landes so perceptibly remarked in his monumental book *The Wealth and Poverty of Nations*, “The weakness of autocracy is in the human raw material. Fortunately.” So, with the collapse of communism in Europe in the 1989–1991 period, the triumph of Millsian “conspiring circumstances” was complete.

Two cases of transition – or transitions, economic but also or even in the first place political – should be considered in this story in somewhat greater detail: the post-communist transition in Europe and the transitions in the two most populous countries of the world: China and India. They will be dealt with in that order.

Post-communist transition in East-Central and Eastern Europe was certainly the most difficult shift toward the greater role of markets and greater openness. Clearly, it was more difficult to build the capitalist market system from scratch, where none had existed after the imposition of communism, than to remove (numerous) distortions imposed upon the still existing, even if severely constrained, market system in countries of the South. Apart from institution-building, both stabilization and liberalization were more difficult. Moreover, the extent of privatization was infinitely greater as well.

What helped in the eight countries stretching from Estonia in the North-East to Slovenia in the South-West were common civilizational, that is Western, roots. Elites and societies in these countries regarded themselves as a part of the West from which they had been forcibly detached by the Soviets. Thus, political system change went there very smoothly. Nowhere else, in the former East or in the South were informal institutions so conducive to such change (maybe in Croatia for the same reasons).

So, the foregoing eight countries (Estonia, Latvia, Lithuania, Poland, Slovakia, Czech Republic, Hungary and Slovenia) became the transition leaders and at some point anchored in Western organizations like NATO and the European Union. The remaining, more numerous post-communist countries fared worse – at times much worse. The reasons, in my opinion, expressed *inter alia* in my 2004 Routledge book *Private Sector after Communism* (written jointly with Vladimir Benacek and Mihaly Laky), are not only in the less consistently pursued transition strategy with respect to stabilization, liberalization, and privatization, but also a much less conducive civilizational environment. The area of Eastern Christendom, for example, never enjoyed the similar extent of civic and economic freedoms over the last millennium. Seen from a historical perspective, informal institutions cast a very long shadow.

The classification of the eight above-mentioned countries as examples of success is sometimes contested. Critics point to the fact that these countries never achieved – after systemic change – very high economic growth rates like, for example, the four “little dragons” in the past or reformed communist China at present. The answer is, however, that they could not for the following reasons:

- They had to shed the deadwood communist system’s production for production’s sake, that is to cut production unneeded under the normal market circumstances. It was primarily, but not exclusively, production of heavy industry.
- Next, gravitational model of foreign trade explains well the rapid expansion of these countries’ exports to Western Europe. What it cannot explain, and predict, is the fact that Western Europe is a slowly – and increasingly slowly – growing part of the world. In spite of an undoubtedly high rate of export expansion to the West and successful incoming foreign direct investment therefrom, the slow growth of the major trade partner matters (and will matter even more).

- Finally, shedding the deadwood of unneeded output resulted in sharp GDP declines in the early transition affecting the average economic growth rate in the longer run. Besides, it is an intellectual error, often committed by those who point to China as a country that succeeded in economic transition without GDP decline, to expect that the East-Central European economies could do the same, but failed in that task. These are middle-developed economies, while the Chinese pattern is possible only at a low level of economic development. Only at that low level there is still a labor force available in agriculture that is willing to work in new, efficient labor-intensive industries emerging in transition to the market. It is increases in their output that may balance the decreases in output in the distorted heavy, capital-intensive industries, dominant at the middle level of economic development. There was no such labor in existence in any significant quantities in the middle developed economies of East-Central Europe.

CHINA AND INDIA

Another set of confusing ideas surrounds China and India. In their case, however, it is more an ideological than purely economic theorizing-based confusion. Some West-haters (capitalism-haters, multinational business-haters, you name it) advertise the thesis on the emerging “Beijing consensus” model allegedly better than a classical Western economic thinking-based “Washington consensus” model.

The reality is – as it often happens – different. Just like with East Asian “little dragons” whose successes were first ignored, then ascribed to luck and special circumstances, and later adopted as another version of interventionism, the Chinese institutional shift to the greater role of markets internally and greater economic openness externally, has also been adopted as another version of interventionism. Although the old Marxian slogan “Proletarians of all countries, unite!” has long been dead, another leftist slogan “Tinkerers of all strands of economic thinking, unite (in negating the success of markets)” is still alive. And increasingly kicking in recent years.

To summarize, the China-India story, whatever economic successes they achieved so far in terms of economic growth and in lifting hundreds of millions from poverty to near-middle class level, has been achieved thanks to the (still far from complete) capitalist market. By contrast, whatever problems they continue to be beset with are largely due to the impact either of the recent past (communist or bureaucratic planning past), or the impact of the more distant past of despotic, stagnant respective empires.

AUSTRIAN BUSINESS CYCLE THEORY WITHOUT UNREALISTIC CONSTRUCTS

DANIIL GORBATENKO¹

ABSTRACT

Austrian Business Cycle Theory (ABCT) purports to explain recurrent clusters of unsustainably lengthy investment projects following episodes of credit expansion by central banks. In this paper, we show that doing justice to the insights behind ABCT requires a disequilibrium-based restatement of the theory and abandonment of unrealistic equilibrium and equilibrium-related constructs. We also sketch the contours of the explanation of the genuine errors that need to be made to set the stage for the unsustainable boom described by ABCT.

KEYWORDS

Austrian Business Cycle Theory, equilibrium, disequilibrium

1 PhD candidate at the Department of Economics and Management (Economie et Gestion) at the Aix-Marseille University; email: daniilgor2004@gmail.com. The author would like to thank P. Garelo, Y. Maouchi, and an anonymous reviewer for helpful comments. The usual caveat applies.

1. INTRODUCTION: THE INTUITIONS BEHIND THE AUSTRIAN BUSINESS CYCLE THEORY AND ITS CURRENT STATE OF DEVELOPMENT

Austrian Business Cycle Theory (ABCT) was developed in an attempt to account for a recurrent pattern of events (normally referred to as the business cycle) under which the boom stage during which the economic activity was high was followed by a period of economic decline. The economists of the British Currency School who were in many ways precursors of the economists of the Austrian School also noticed that this cyclical pattern was preceded by attempts of the institutions controlling the money supply (in the British Currency School's case, the Bank of England) to expand the money supply. In this view, the expansion of the money supply proceeded to create excess credit beyond the supply of voluntary savings available in the economy. The Currency School's economists referred to this phenomenon as "forced saving".²

The economists of the Austrian School have deepened these intuitions. The understanding of the heterogeneity of capital and the explicit treatment of production as happening in time allowed them to conjecture that the boom phase of the cycle consisted in the deviation of the time structure of production from the time structure of consumer demands. The deviation consisted in a cluster of investment projects being undertaken that later turned out to be unsustainable.

The theory which emerged from these intuitions was, however, rather aggregative. Instead of focusing on the cluster of particular investment projects and the loans by which they are financed, Ludwig von Mises (2008 (1949)), F.A. Hayek (2008 (1931)) and those who followed in their steps tried to use several problematic constructs.³

In general terms, the classical version of ABCT can be summarized as follows. In the normal circumstances in which voluntary savings are the only source of credit, the equilibrium loan market interest rate (gross market interest rate) stays equal to the natural (originary) rate of interest. The latter is defined as the equilibrium interest rate which would have been established if the lenders lent investors not money but intermediate goods directly. In case of a credit expansion by the central bank, since the loan market interest rate falls below the natural interest rate, it becomes illusorily profitable for the producers to make the production structure of the economy more roundabout or to increase the number of stages of production.⁴

2 For a detailed discussion of the British Currency School's view see Mises (2008 (1949), 568–73).

3 For notable more recent versions of ABCT, which in our view employ reasoning similar to that of the early accounts of Hayek and Mises, see Rothbard (1996 (1978)), Garrison (2000), Murphy (2011), Salerno (2012).

4 Some economists of the Austrian School believe that credit expansion may also be caused by a concerted action of private fractional-reserve banks in an economy without a central bank. For the reasons provided by Selgin (1988, 80–82), we are sceptical of such a possibility. Nevertheless, even if decentralized credit expansion is possible, it does not impinge on the arguments made in this paper. They would apply equally well to such a case.

However, since consumers have not, in fact, changed their preferences in favor of abstaining from some consumption, the lengthening of the production structure turns out to be unsustainable when they attempt to readjust the production structure back in line with their demands. Because the stock of scarce intermediate goods⁵ has not increased to enable the coexistence of the non-decreased consumption and elongated production structure, something has to give.

The banking system can keep the troubled investment projects afloat for some time by accelerating the credit expansion and lending to them at even lower rates. But at some point, the central bank becomes concerned with the inflation, which tends to increase because of the leakage of the additionally created money into the markets for consumer goods. It tightens the monetary policy stance, and banks cannot continue to accelerate lending to the unsustainably long production processes. As a result, those processes have to be frozen or abandoned.

Finally, because of the relative specificity of certain capital goods used in the troubled projects, it may be difficult or even impossible to reallocate them to the uses which better serve consumers. Thus, part of the capital of the economy becomes trapped or wasted, necessitating temporary economic decline.

There are some intuitions reflected in the classical version of ABCT which are promising. The first of them is the idea that during the boom stage, there is a deviation in some sense of the time structure (or pattern) of production from the time structure (or pattern) of the consumer demands. Another intuition that is worth mentioning explicitly is that there is a real physical scarcity of certain intermediate goods that is exacerbated during the boom and ultimately renders certain production processes unsustainable.

At the same time, the particular way in which those intuitions are formulated is highly problematic. The first major problem which we will deal with in this section is that certain mental constructs on which they are based are highly problematic.

The notion of aggregate roundaboutness measured by the number of stages of production implies that there must be an uncontroversial way of assigning a particular capital good to a particular stage of production at which it is produced. However, a simple observation shows that it is impossible. It has usually been conceived that extraction of the raw materials represents the earliest, or highest, stage of production. However, the extraction of most raw materials requires the use of certain capital goods that need to be produced before it takes place. It is not difficult to see a regress to an unclear end which is, however, obviously not very relevant to the issues of the modern business cycle.⁶

5 In this paper, by “intermediate goods” we will mean all goods which are used in the production of final consumer goods. Intermediate goods are traditionally classified into capital goods and original factors of production (raw materials). While this classification is not without problems (for instance, it is difficult to classify electricity under it), in our view they are immaterial for the arguments put forward here.

6 This point was first raised in Young (2011, 27).

Thus, in practice, in order to speak of aggregate roundaboutness, one has to try to define the measure in terms of which a more roundabout production structure differs in the aggregate from a less roundabout one. This measure must in some sense be physical because, as we mentioned in section 1, one of the key intuitions behind ABCT is that a real physical scarcity of intermediate goods is involved in the business cycle.

Hayek tried to solve this problem by invoking the ratio between the physical volumes of the output of intermediate and consumer goods. That he was talking about physical amounts is obvious from the following passage:

Probably the simplest method of transforming the picture of the continuous process into a picture of what happens in a given period is to make cross sections through our first figure at intervals corresponding to the periods chosen, and to imagine observers being posted at each of these cross cuts who watch and *note down the amount of goods*⁷ flowing by.⁸

The figure referred to in the quote is the famous Hayekian triangle. For Hayek the changes in the structure in production involved *inter alia* the changes in the proportion between the physical volumes of intermediate goods vs consumer goods for a given period which tend to mirror the changes in the proportion of money streams flowing into their purchases.

In such a case, the proportion of money spent for consumers' goods and money spent for intermediate products is equal to the proportion between the total demand for consumers' goods and the total demand for the intermediate products necessary for their continuous production; and this, in turn, must correspond, in a state of equilibrium, to the proportion between the output of consumers' goods during a period of time and the output of intermediate products of all earlier stages during the same period.⁹

The boom phase of the business cycle, in this view, consists in the change of such proportion during the boom period which is not compatible with consumers' preferences. At the turning point of the cycle, consumers act to reverse that proportion making certain investment projects unsustainable.

The main problem with this reasoning is that it essentially implies that it is possible to strictly separate all the original factors of production in a given period into those that are embodied in consumer goods and those that are embodied in intermediate goods. The lat-

7 Emphasis mine - D. G.

8 Hayek (2008 (1931), 232).

9 Ibid. P. 234.

ter must not be used to produce consumer goods within the period.¹⁰ Only in this case it would make sense to talk about the changed physical proportions between producer and consumer goods for a period which are equal to the changes in the proportion of the demands for producer and consumer goods. The changes in the structure of production would then mean merely redistribution of the physical quantity from consumer to producer goods or vice versa.

However, certain original factors are, in fact, not fully incorporated into the consumer or producer goods which they are used to manufacture. Consider the example of the production of meat. Although estimates vary, a very large amount of water is needed to produce a kilogram of meat. Yet, most of it will not be incorporated into the meat.

In addition to this, even if relative changes in demand for producer goods and consumer goods only resulted in a changed identifiable physical proportion between the consumer and producer goods produced, it would still not follow that the latter would be equal to the new proportion between the money spent on producer and consumer goods. The value of an intermediate good does not stem from its physical quantity but from the expected value of the consumer goods which it is used to produce. There is, thus, nothing that would preclude a scenario where a change in the structure of production would involve the production of a highly massive stock of producer goods instead of certain consumer goods but a relatively much smaller redistribution of monetary streams.

The notion of a single loan market interest rate is also of dubious utility. A bank demands a certain rate of interest on the loan on the basis of other available alternatives for lending. Since the alternatives are different and there probably are never enough alternatives of the same attractiveness to accommodate the whole supply of credit, there will always be different interest rates on loans which will never converge to a single rate.¹¹

The construct of the natural (originary) interest rate at least in the business cycle context is even more dubious. First, it is not clear what an interest rate could even mean if lending happened directly in intermediate goods. Since those goods would be used in the production of other goods, it would have been impossible to repay the loan in units of the originally lent goods. Thus, the notion of an interest rate ceases to make sense. Another problem first noted by P. Straffa (1932) and discussed in some detail by R. Murphy (2011) is that even if

10 Otherwise, the distinction that Hayek was making would be meaningless because the essential idea is that during a period certain physical quantity can be strictly classified into a part that contributes to the production of final consumer goods and the one which does not. Thus, Hayek's distinction should have been not between the produced physical volumes of consumer and intermediate goods but between, on one hand, the physical volumes of consumer goods and the intermediate goods used in their production but not embodied in them, and, on the other hand, all other intermediate goods.

11 We did not even consider the risk premium as a component of actual interest rates which will vary in accordance with the attitude to risk of particular lenders, as well as with the riskiness of particular projects to be financed.

the notion of the interest rate as applied to lending in the units of intermediate goods made sense, each lent intermediate good would have its own natural rate of interest.¹²

At a more basic level, the use of problematic aggregative constructs which we have discussed seems to arise from a more fundamental problem, i.e. the idea that the theory of the business cycle must be formulated in terms of tendency toward a general equilibrium.

Hayek was especially explicit in his conviction that only equilibrium-based reasoning can account for the phenomenon of the business cycle:

On the contrary, it is my conviction that if we want to explain economic phenomena at all, we have no means available but to build on the foundations given by the concept of a tendency toward an equilibrium. For it is this concept alone which permits us to explain fundamental phenomena like the determination of prices and incomes, an understanding of which is essential to any explanation of fluctuation of production.¹³

Mises's account of ABCT in *Human Action* is also dependent on an equilibrium construct, although somewhat different from the Walrasian one employed by Hayek, i.e. the final state of rest.¹⁴ However, the differences between those two equilibrium constructs are not essential for the purposes of the argument made here because both of them imply that all the factors operative at a certain moment in time will tend to be reflected in the prices to which they are relevant. One type of factor that affects prices is the various plans. In equilibrium-based analysis, all plans must be mutually adjusted to become compatible in order for equilibrium to be achieved.

The connection between general equilibrium reasoning and aggregative constructs arises because general equilibrium analysis, in order to be tractable, requires certain aggregative factors that change (like the gross market interest rate) and certain aggregative outcomes that result (like the total roundaboutness of the structure of production). If we can show that equilibrium-based reasoning can and should be rejected from the account of the business cycle, we can also show that the aforementioned aggregative constructs are unnecessary for ABCT.

12 Note, however, that Murphy is only able to provide an example which makes sense because he focused on the case of a hypothetical economy with two consumer goods that can be exchanged *inter alia* for claims to future amounts of those same consumer goods. It is possible to set the interest rate in terms of a units of a lent good if one borrows apples and produces apples, but it is difficult to see how an interest rate can be set in units of, for instance, oil if one is producing synthetic clothes.

13 Hayek (2008 (1931), 225).

14 See Mises's (2008 (1949)) descriptions of price adjustments in chapter XX. They are given in terms of movement towards the final state of rest.

2. A DISAGGREGATED DESCRIPTION OF THE BUSINESS CYCLE PATTERN

The fundamental point behind the reasoning in this paper is that it is indeed possible to achieve the goal mentioned in the end of the previous section if the theoretical description centers on the primary phenomenon that requires explanation, namely, the clusters of unsustainable long-term investment projects that are started in the beginning of the boom stage but become unsustainable at the turning point of the cycle (further referred to as the “excessively long projects”). The key feature of these projects is that they do not result (directly or indirectly)¹⁵ in the production of consumer goods during the boom period. These projects are not necessarily intertwined directly with the whole economy, thus an attempt to explain them does not require making formal statements about the economy as a whole and certain aggregate elements or properties of it.

A proper emphasis on such clusters of projects means that it must be immediately clear from the resulting version of the theory how those projects are mistakenly undertaken and then fail. If no bridge can be made from a version of ABCT to those projects, then such a version is necessarily inadequate for explaining what it is called upon to explain.

Such an emphasis leads us to the crucial question of why the excessively long projects may fail. In this regard, the essential observation is that the excessively long investment projects fail before they start delivering final consumer goods to the market. This observation suggests three possible types of economic causes of their failure:

1. Their originators realize that they were mistaken in their assessment of the future demand for the relevant consumer goods.
2. Certain innovations make it more profitable to use the resources involved in the relevant projects to be employed elsewhere.
3. There is an unexpected increase in costs of the projects.

The first possibility seems to be precluded by the fact that we are dealing with a cluster of investment projects. If we take into the account the coordinative properties of the market process, they seem to cast doubt on the possibility of systematic spontaneous entrepreneurial mistakes of the kind that are necessary.¹⁶

The second possibility does not square well with the fact that the failure of the excessively long investment projects marks the beginning of the bust stage of the business cycle. Innovations should tend to result in economic growth, not in decline. Thus, we are left with the third alternative, i.e. that the unsustainability of the excessively long projects results from an unforeseen increase in the costs of their implementation.

15 An investment project that does not itself lead to production of consumer goods in a given period may still create such intermediate goods that are part of the chain resulting in the production of consumer goods in such a period.

16 For a more detailed discussion of why the idea of spontaneous concentrations of entrepreneurial forecasting errors is problematic, see Rothbard (1996 (1978), 70–73).

We will be assuming here that the costs of investment projects are well approximated by the expenditures on labor and intermediate goods necessary to complete them. While ABCT can in principle be extended to the analysis of the changes in the pattern of allocation of labor, we will limit our analysis to the intermediate goods for the sake of keeping the paper sufficiently tractable.

It is clear from the intuitions underlying ABCT that the unexpected growth of costs that hits the excessively long projects is of an endogenous nature. Otherwise, the whole idea of the business cycle would have been unfounded. In other words, there must be some process internal to the market which aggravates the always existing scarcity of intermediate goods which somehow arises from the mistaken decisions of the economic agents competing with each other for those goods.

The process internal to the market is competition. Thus, the excessively long projects must compete with some other projects in an unsustainable way for certain resources (the “contested goods”). Unlike the former, the latter projects (hereinafter, the “closer-to-production projects”) must result (directly or indirectly)¹⁷ in the production of final consumer goods at some point during the boom stage. This has to be the case because, otherwise, there would be no role for consumer preferences in bringing about the reversion of the wrong pattern of resource allocation.

The closer-to-production projects do not necessarily have to take less time to deliver final consumer goods than the excessively long projects in absolute terms. Rather, at the moment at which the latter projects are started, the former must be closer to the production of the final consumer goods. This allows us to avoid the thorny issues of talking about stages of production and the increase in their number.

The final building block that we need for our description derives from the character of business planning undertaken in time. In any production, producers cannot just base their planning on wild guesses about what goods the consumers will prefer to buy. They must ground their plans in the relevant experience from the preceding period or periods (the “reference period(s)”). Thus, the originators of the closer-to-production projects must relate their plans to the reference period(s). For the sake of simplicity, we will assume in this paper that no change in consumer preferences with regard to the consumer goods produced by the closer-to-production projects happens in the boom stage compared to the reference period(s) of the originators of the closer-to-production projects.¹⁸ We also assume that there is no technological change with respect to the production of the relevant consumer goods,

17 See footnote 15.

18 It is beyond the scope of this paper to give a detailed descriptions of this process, but it is not necessary for its purpose. However complex the use of past experience by entrepreneurs might be, what matters for the theory contained here is that there is a certain picture of the particular market patterns on which the business planning by the various agents discussed in this paper is based. Those agents might believe that the boom period will be exactly the same as the previous period of the same calendar duration, or they might base their calculations on a certain generalization of several such reference periods. What the monetary expansion must bring about is a deviation of the relevant market process from the expected pattern in the relevant market.

and that no increase in their production or production of their substitutes happens in the boom period.

The originators of investment projects must also look into the past to forecast the prices for the intermediate goods which they will use, in this case the contested goods. Here we assume that in the boom stage of the cycle nothing changes with regard to the demand and supply of the contested goods, except for the unforeseen additional demand from the excessively long projects.

There is also a third class of entrepreneurs which have been ignored in the previous versions of ABCT whose decisions are, however, indispensable for producing the boom pattern. These are the suppliers of the contested goods. They also must ground their planning for the boom period on past reference periods. We will discuss their role in more detail in the next section. Finally, the same kind of reasoning applies *mutatis mutandis* to the banks that allocate the excess credit whose creation is enabled by the central bank.

The preparatory steps of the argument which we made above allow us to zoom in on the crucial but vague intuition underlying ABCT in a straightforward way. The idea is that during the boom stage of the business cycle the time structure of production deviates in some sense from the time structure of consumer demands. This deviation may now be said to consist in a cluster of excessively long investment projects temporarily diverting part of the contested goods from the closer-to-production projects not in accordance with consumer preferences.

To clarify this general statement, at some point during the boom stage of the cycle, the originators of the closer-to-production projects discover that they are unable to supply their customers with as much of certain consumer goods as they were planning to. This happens because part of the contested goods has (unexpectedly for the originators of the closer-to-production projects) been consumed by the excessively long projects.

At the same time, since the preferences of the relevant consumers with regard to the consumer goods that are produced under the closer-to-production projects have not changed compared to the reference period(s), more money is now chasing a lower quantity of goods compared to the reference period(s).

The originators of the closer-to-production projects learn this by means of price signals. They observe that they are able to demand higher prices for the consumer goods that they produce than in the reference period(s), and they know that they are supplying fewer of them in the current period. To the extent that it is still profitable for them to increase production of the relevant consumer goods, despite the need to compete for the contested goods with the excessively long projects, they will tend to grasp such a profit opportunity. This increased competition for the contested goods will tend to drive up their prices.

But how could this turn of events be fatal for the excessively long investment projects? The only possible answer is that they must be tailored to the prices of the contested goods at which their originators are initially able to acquire them. They must not necessarily become unsustainable with any increase in the prices of the contested goods which go into their calculation. It is sufficient that the potential increase in their prices for which the calculations of the projects provide is lower than the actual increase arising through the process described above.

This condition for the unsustainability of the excessively long investment projects is quite plausible. After all, at the stage of planning the projects, the entrepreneurs who consider undertaking them must make use of the then current prices of the contested goods. Based on the reference period(s) they will likely make provisions for those prices fluctuating within a certain range. However, this safety range may well turn out to be not sufficient upward to absorb the aforementioned increase in the prices of the contested goods.

At the same time, the originators of the excessively long projects must initially be able to purchase part of the contested goods at prices which are within the aforementioned safe range. Otherwise, the excessively long projects would not be undertaken. We will address the question of how the formation of this first set of prices of the contested goods may happen in section 4.¹⁹

Finally, as in the classical versions of the theory, the banking system may try to help the excessively long projects by lending to them at even lower interest rates than in the beginning. However, this strategy will tend to be self-defeating because at this point it will only tend to result in the increases in the prices of the contested goods, offsetting the effect of the lowered interest rates. At some point, banks will realize that it is not profitable to finance the excessively long projects.

3. THE COMPATIBILITY WITH EQUILIBRIUM CONSTRUCTS

At this point, we can return to the problem of compatibility of the equilibrium-based formulations of ABCT with the properly focused general description of the boom stage of the cycle which we have just made. We have mentioned in section 1 that the hitherto published prominent versions of ABCT have all been based on equilibrium reasoning. However, we argued in the preceding section that any version of ABCT must be compatible with the primary phenomenon it purports to explain. We also sketched out a theoretical description of the boom stage of the business cycle directly in terms of that phenomenon, namely in terms of a cluster of excessively long investment projects.

The upshot of the above is that if equilibrium-based reasoning is incompatible with this description, it would strongly suggest that it is incapable of accounting for the subject matter of ABCT and should thus be rejected.

Clearly, the description that we made in the preceding section is incompatible with a single act of equilibration. After all, under the pattern that we just described, at first, the originators of the excessively long investment projects are able to acquire part of the contested goods

19 It does not seem to matter for the validity of the argument of this paper whether the originators of the excessively long projects are able to buy part of the contested goods at safe prices only once or several times. Thus, we can, for the sake of simplicity refer to those safe prices as the first set of prices of contested goods and to the prices which reveal the unsustainability of the excessively long projects as the second set of such prices.

that they require at one set of prices which are lower than the prices they will face when it is discovered that they have diverted the contested goods from the closer-to-production projects in an unsustainable way. Two sets of prices require two instances of equilibration.

In light of this, the question becomes whether the business cycle pattern can be accounted for by means of an inter-temporal equilibrium construct. This question also needs to be answered in the negative because of the logical structure of the pattern that needs to be explained. We need to recall that in the most abstract terms, what we deal with are two types of projects which are ultimately incompatible in the sense that not enough contested goods are available to complete both of them at the prices to which the excessively long investment projects are tailored. Meanwhile, both types of projects are for some time being implemented simultaneously until their mutual incompatibility is discovered when it is reflected in the prices of the contested goods.

In other words, the formation of the second set of prices²⁰ of the contested goods, which reveal the unsustainability of the excessively long projects, is not brought about by an exogenous change in which case the pattern would be consistent with the inter-temporal equilibrium. Rather, its emergence is an endogenous result of those projects' being incompatible from the start but being undertaken for some time simultaneously. In a pattern consistent with inter-temporal equilibrium, the two incompatible types of projects would have to undergo adjustment before starting to be implemented.

One possible objection remains to be answered here. It may be argued that inter-temporal equilibrium is a benchmark with which the course of events during the business cycle should be contrasted. In other words, one could say that in the case of an increase of voluntary savings various investment plans tend to be equilibrated over time while in the case of the artificial boom they become dis-coordinated. Here it must be noted that we do not object to the use of the notion of equilibrium in this way. However, this usage is fundamentally different from that which we have discussed above. It leaves space for genuine errors, and does not attempt to impose a certain straitjacket on reality. Perhaps, for the purpose of avoiding confusion it is better to speak not about equilibrium but the degree of coordination among plans which is and can never be perfect.

4. HOW ERRORS ARE MADE MORE PROBABLE BY CREDIT EXPANSION

In the preceding sections we dealt with the way the business cycle needs to be described. Now that we have established that the business cycle pattern can only be consistently described in disequilibrium terms, we need to explain how the genuine entrepreneurial errors which are possible in a disequilibrium context may plausibly come about. We use the words “may plausibly come about” deliberately because abandonment of the equilibrium frame-

²⁰ See footnote 19.

work necessarily leads to abandonment of an apodictically certain link between the credit expansion and the cluster of unsustainable investment projects. This may sound like an unacceptable theoretical price to pay but we will try to demonstrate further in this section that there is a silver lining to this sacrifice.

However, before we start explaining how the malinvestment pattern may plausibly happen because of credit expansion, we must ask another question. How do markets achieve the relative coordination between the inter-temporal structure of production and the inter-temporal structure of consumer preferences in the case of a somewhat similar situation not involving credit expansion? We are talking here about a scenario with an increase in the supply of voluntary savings that are made available to the banking system. Echoing the famous comment by Hayek on Keynes's theory of macroeconomic instability, it is essential to explain how the coordination mechanism works in the normal case in order to understand how it might fail in the case of credit expansion.

In other words, the question is why monetary savings made by consumers will tend to be channeled into the right projects and why credit expansion will tend to cause an inter-temporal disconnection between investment projects and consumer preferences. To answer these questions, we need to consider the elementary mechanics of a bank investment loan.

In order to draw down an investment loan, the entrepreneur must believe that the project that will be financed with this loan will provide a return which will exceed the amount of the money she will owe to the bank by at least a certain amount. This implies that the decision to draw down a loan will only be made if the entrepreneur expects a certain minimum level of profitability of the relevant project.

The availability of the additional money to lend for the banks as a result of an increase of savings does not in itself change the expected profitability of any investment project. What changes the expected profitability of some projects is the fact that the corollary decrease in the demand for consumer goods lowers the prices of some intermediate goods that can be used in those projects. To the extent that entrepreneurs recognize the increased expected profitability of some projects, this circumstance tends to cause the demand for investment credit on their part to rise and to allow the banks to loan out the additional money they received from the savers.

But why will the additional amount of money arising from the savers' making savings through the banks tend to be allocated to the projects which have become more profitable in the way described above? The answer has to do with our description of business planning in time in section 2. Banks also have to ground their lending plans in the reference period(s). If there is no significant change in the boom period in the other projects that they would finance compared to the reference periods, banks will not tend to allocate the additional money to them if there are investment projects which are now more profitable even without any decrease in the interest rates. Thus, the additional money will tend to be allocated to the projects which have become more profitable as a result of the increase in voluntary saving without any connection to interest rates.

It is important to note here that the average interest rate in the loan market may not even fall as a result, contrary to what is usually claimed. This is not a violation of the basic sup-

ply and demand reasoning because the fact that some projects became profitable compared to the reference period(s) results in an increased supply of loanable funds being met by an increased demand.

Thus, the crucial inter-temporal coordinative role is played by the prices of the relevant intermediate goods. In the classical accounts of ABCT, their role was not sufficiently recognized and excessive attention was paid to the aggregates of interest rates.

Now that we have looked into the way the coordination works in the case of the additional credit arising from an increase in savings, we can start addressing the crucial question of why a credit expansion may plausibly result in the discoordination described by ABCT. Intuitively, we need to demonstrate that something important changes vis-à-vis the normal credit creation scenario.

This important change consists of there being no decrease in the prices of intermediary goods because the consumers have not reduced their consumption. Thus, a crucial pillar of the coordinative mechanism is missing from the start. The banks that want to create credit out of thin air to allocate to investment projects do not face offers from entrepreneurs whose projects have become profitable because of the decreases in costs. Thus, banks have to induce the potential borrowers to borrow from them by lowering interest rates on the loans involving the excess credit and making certain projects look more profitable.

Two questions remain to be answered here. First, why are such loans significantly more likely to be allocated to the excessively long investment projects? The answer has to do with the reasoning of the preceding section. If the investment projects which receive the excess credit are not excessively long, their competition with the other projects for the relevant intermediate goods will be apparent from the very start from the increases in prices of such goods. Those price increases will tend to negatively compensate the increased profitability due to the lowered interest rates. In practice, the loans of the excessive credit to such projects will probably mostly not even be made because the suppliers of the relevant intermediate goods will be in position to increase the quoted prices for them from the very start. This will tend to reveal the unsustainability of the projects which are not excessively long but have become profitable due to the lowered interest rates at the stage of negotiating the terms of the loans. Thus, to the extent the excess credit created by the central bank is actually loaned out, it will tend to be allocated to the excessively long projects.²¹

The second question is why banks, the entrepreneurs who borrow the excess credit for the excessively long investment projects and, as we mentioned in section 1, the suppliers of the contested goods will tend to fail to see that they are making errors in their commercial decisions. In other words, we need to describe in what conditions credit expansion makes likely the genuine errors that we mentioned in section 3.

21 Since the market process (including in the banking sphere) works only imperfectly, it is theoretically possible that the projects which banks are prepared to lend to when they are able to create excess credit have actually been sufficiently profitable before, and that they did not receive loans by mistake. Thus, excess credit may theoretically be allocated to good projects. However, the more excess credit is created, the more likely it is that it will go into excessively long projects.

The first problem that we face here is that, in principle, entrepreneurs might be aware of ABCT and might thus avoid errors by not borrowing the money that arises from credit expansion. The solution to the problem lies in realizing that an entrepreneur cannot be certain of the fact that either her project is going to be financed by means of credit expansion or that her particular project is going to be unsustainable. Here the abandonment of the necessary connection in ABCT between credit expansion and the cluster of malinvestments provides the very silver lining we mentioned in the beginning of this section. Since there is no such necessary connection that could be understood by entrepreneurs, even if they knew about ABCT, they would still face a choice between foregoing the potential profit from the investment project financed by means of credit expansion and risking failure of such a project due to the mechanism described by ABCT. Thus, we can say that even some entrepreneurs knowledgeable about ABCT may be lured towards undertaking unsustainable investment projects. It also seems to follow that the larger the scope of credit expansion, the higher the risk becomes that some of the entrepreneurs will choose to try their luck because with the increasing amount of credit, the interest rates will be set increasingly lower, thus reducing the total amount of debt from the loans.

Of course, the vast majority of entrepreneurs never were, are and probably will be aware of ABCT in any meaningful sense, but this counterargument needs to be addressed in light of the popularity in modern economics of certain interpretations of the rational expectations hypothesis.

The same reasoning applies to banks with the modification that the risks they bear in this situation are even less significant because for them the alternative consists not in some other use of their resources but in not using some credit at all.

The final question which remains to be answered is why the sellers of the contested goods who must initially sell them to the entrepreneurs undertaking the excessively long projects at prices that are too low, i.e. do not reflect the changed structure of demands for these goods compared to the reference period(s), may make this error. Here, although the sellers of the contested goods might be able to realize that the demand for them has increased compared to the reference period(s), in certain plausible conditions it may be difficult for them to do so. For example, if the lengthier investment projects are not very large, and (or) if they are not concentrated in a narrow industry, and (or) if the sellers are not very concentrated. Because of the considerations of brevity, we may not plunge ourselves deeply into the discussion of various combinations of such and similar conditions, but it is clear from this short analysis that a situation where the sellers of the contested goods fail to timely notice the increase in demand is fairly plausible and does not require any unrealistic assumptions. Moreover, the larger the scope and extent of the credit expansion, the more likely are some sellers of the contested goods to commit price-setting errors.

In other words, the key feature of the additional investment demand for intermediate goods that can be created by credit expansion, is that it ultimately arises arbitrarily from the top without any connection to the underlying market process, and thus to the relevant knowledge of the suppliers of intermediate goods. Thus, where such arbitrarily created change is not momentous, it may not be noticed on time and may create discoordination.

On a final note, the original mistakes of the suppliers of the contested goods may trigger further mistakes which are usually not thought to be the subject of ABCT. If the producers of the contested goods, having realized that the demand for their products for the relevant period is going to be higher than they had expected, attempt to accommodate it by investing into increases in production, they will still have to charge higher prices for their products because they bear additional costs. However, as we established, these increased prices are unacceptable for the originators of the excessively long projects. Thus, at the end of the day, it is not just the originators of the excessively lengthy investment projects who may make malinvestments but also the producers of the contested goods. It is even possible that errors can be made further down the chain by the suppliers of the producers of the contested goods and even their suppliers, but a detailed discussion of these issues is beyond the scope of this paper.

5. CONCLUDING REMARKS. POTENTIAL APPLICATIONS

We have demonstrated in the preceding pages of this paper that the crucial insights of ABCT neither need, nor can be formulated by means of an equilibrium construct. Attempting to do this leads to the paradox under which the theory precludes the phenomena, which it is supposed to explain, from happening. We also attempted to formulate in general terms a genuinely disequilibrium-based account of ABCT.

This paper is only sketching the contours of the theory; thus, it only deals with an outline of the necessary conditions for the Austrian business cycle pattern to take place. But even such cursory analysis has, in our view, important implications that go beyond the issues discussed here.

One of these important points is that even the economists of the Austrian School, who are generally much more realistic in their treatment of economic phenomena than their mainstream colleagues, do not have at this point a satisfactory theory of price formation and changes. Perhaps, such theory cannot be formulated in a satisfactory manner at all due to the inherent complexity of the issues involved, but it seems that these issues warrant revisiting by the modern economists of the Austrian school in order to make the theories more tractable and convincing.

The second major implication is that in some cases we need to abandon the apodictic character of the theory in order to preserve its relevance to the reality under consideration. This may be a hard pill to swallow for some Austrian economists but the idea is not new. The ideas elaborated on in this paper are in line with the insightful criticisms of ABCT by Hülsmann (1998) and Wagner (1999) who have both demonstrated that the insights contained in ABCT are not congruent with a logically necessary connection between credit expansion and entrepreneurial errors leading to the boom and bust.

Finally, the realistic approach to ABCT simplifies the job of those researchers who would attempt to provide an empirical illustration of the theory on the basis of the histori-

cal instances of the business cycle. It gives them real-world reference points that they might try to locate and theoretically connect in the events of the relevant episodes, such as the cluster of investment projects, the contested intermediary goods and their price dynamics, etc., instead of merely looking at certain statistical aggregates.

In other words, the concepts of credit expansion, contested goods, excessively long projects and closer-to-production projects allow for genuine historical, rather than merely aggregative econometric analysis of the past business cycle episodes, quite in line with Mises's notion of applying theory to history.²² It is possible to try to find in the historical episode in question the referents of the aforementioned straightforwardly applicable concepts. If such referents are found, a conclusion can be made that a given historical episode is an instance of the pattern described by ABCT. Otherwise, the economic historians studying a particular episode will have to conclude that some other theory is needed to explain.

As a brief example of how such analysis may proceed, it can be noted that with regard to the most recent episode of the business cycle in the US, which partly consisted of a bubble in the market for new single-family houses, several Austrians have claimed that it is well-explained by ABCT.²³ However, the reasoning contained here casts a serious doubt on this conclusion for the simple reason that investment projects for building the single-family houses in question just do not take a sufficient amount of time.²⁴

REFERENCES

- Boettke, P., Horwitz, S. (2009) *The House That Uncle Sam Built: The Untold Story of the Great Recession of 2008*. Foundation for Economic Education. Accessed September 4, 2014. <http://www.fee.org/library/detail/the-house-that-uncle-sam-built>.
- Garrison, R. W. (2000) *Time and Money: The Macroeconomics of Capital Structure*. Routledge; Reissue edition.
- Hayek, F. A. (2008 (1931)) *Prices and Production (Large Print Edition): And Other Works on Money, the Business cycle, and the Gold Standard*. CreateSpace Independent Publishing Platform.
- Hülsmann, J. (1998) 'Toward a General Theory of Error Cycles', *Quarterly Journal of Austrian Economics* 1 (4): 1-23.
- Mises, L. v. (2008 (1949)) *Human Action: A Treatise on Economics*. Auburn, Alabama: Ludwig von Mises Institute.

22 See Mises (2007 (1957)) for a book-length treatment of these issues.

23 See e. g. Boettke and Horwitz (2009), Salerno (2012).

24 Saniavskaja (2013) estimates that it takes on average 5-6 months to build a single-family house in the US. Clearly, the construction projects with respect to such housing do not qualify as the excessively long projects that should have failed if the recession of 2008 was an instance of the boom and bust process discussed here.

- Mises, L. v. (2007 (1957)) *Theory and History: An Interpretation of Social and Economic Evolution*. Auburn, Alabama: Ludwig von Mises Institute.
- Murphy, R. P. (2011) *Multiple Interest Rates and the Austrian Business Cycle Theory*.
- Rothbard, M. N. (1996 (1978)) 'Economic Depressions: Their Cause and Cure', In *The Austrian Theory of the Trade Cycle and Other Essays*, edited by Richard M. Ebeling. Auburn, Alabama: Ludwig von Mises Institute .
- Salerno, J. T. (2012) 'A Reformulation of the Austrian Business Cycle Theory in Light of the Financial Crisis', *The Quarterly Journal of Austrian Economics* 15 (1): 3-44.
- Selgin, G. (1988) *The Theory of Free Banking*. Boston Way, Lanham, MD: Rowman & Littlefield Publishers, Inc.
- Siniavskaia, N. (2013) *How Long Does It Take to Build a House?* Eye On Housing, October 21. Accessed September 4, 2014. <http://eyeonhousing.org/2013/10/how-long-does-it-take-to-build-a-house/>.
- Straffa, P. (1932) 'Dr. Hayek on Money and Capital', *The Economic Journal* 42 (165): 42-53.
- Wagner, R. (1999) 'Austrian Cycle Theory: Saving the Wheat while Discarding the Chaff', *Review of Austrian Economics* 12: 65-80.
- Young, A. (2011) 'Illustrating the Importance of Austrian Business Cycle Theory: A Reply to Murphy, Barnett, and Block: A Call for Quantitative Study', *Review of Austrian Economics* 25 (2): 19-28.

HETEROGENEOUS CAPITAL AND THE COASEAN FIRM – A CRITIQUE OF SOME RECENT DEVELOPMENTS IN THE AUSTRIAN THEORY OF THE FIRM

IVAN JANKOVIC¹

ABSTRACT

In this paper some recent improvements and refinements of the mainstream Austrian theory of the firm are analyzed (Foss and Klein, 2012; Foss et al, 2007). Those refinements attempt to fuse the old Austrian-Coasean theory of the firm with a sophisticated analysis of entrepreneurship and capital theory, in order to broaden the scope and the application of the original intuition behind the Coasean theory that governance, planning and hierarchical mode of relationship are necessary features of an efficient market system. I criticize this strategy by combining the insights of the Austrian capital theory and the standard neoclassical theory of the firm, to show how the Coasean firm, both in its original and Austrian derivations misrepresents the real reasons and function of the firm in a market society. The main problem of this new synthesis is that it just broadens the old error committed by Mises and Rothbard, treating Coasean theory of the firm as central planning hierarchy to be consistent with Austrian argument against planning, and for economic calculation and entrepreneurship. In my analysis, the ingenious inclusion of the Austrian capital theory into the Coasean framework by Foss and Klein only adds another layer of errors to the original, untenable conception of the firm as a hierarchical and organizational entity by arguing that the firm is needed to handle the problem of heterogeneity of capital goods.

KEYWORDS

Capital theory, firm, Coase, Austrian School.

¹ Doctoral candidate in political science , Simon Fraser University, Vancouver; email: ijankovi@sfu.ca.

1. INTRODUCTION

The conventional treatment of the theory of the firm in the modern literature of the Austrian school combines Mises', Hayek's and Rothbard's emphasis on entrepreneurship and economic calculation with the concept of the firm proposed by Ronald Coase and is widely accepted among the students of economic organization (Foos and Klein, 2012; Foss et al, 2007; Ionadies, 1999). This theory depicts the firm as a central planning and hierarchical entity, supplementing or rather supplanting the market as a coordinating mechanism. Principally, it combines Coase and Mises, i.e. develops a theory of the firm as a commanding authority serving the purpose of "organizing the entrepreneurial judgement" as Foss and Klein like to say. This dominant tradition of organizational thinking within the Austrian school will be called "Austrian Coaseanism" or simply the "mainstream Austrian theory of the firm". A minority of authors writing in the Austrian tradition question this amalgamation of Misesian and Coasean traditions and argue that the firm should be understood as a specific market mechanism (Salin, 2002; Mathews, 1998; Jankovic, 2010).

There are many other aspects related to the Austrian theory of the firm treated in other works, but they are only tangentially related to the main problem of the theory of firm. For example, Dulbeco and Garrouste (1999) propose a unification between the capital theory and theory of entrepreneurship, but here the firm is conceptualized just as a kind of black box with the tools necessary to explain the main problems such as individual knowledge and asset specificity. The theory is a functional one in the same sense the neoclassical theory of the firm is functional, striving to explain the exchange on an open market. Salin (2002), ties the discussion of the firm to the problem of division of labor; Walsh (2009) offers an ingenious Mengerian invisible-hand explanation of the origins of the business firm trying to strike a balance between the notions of spontaneous order and conventional treatments of the firm; Cowen and Parker (1997) explore the problems of reconciling the theory of the firm and the process-character of the market exchange. Garrouste (2002) analyzes the implication of the Hayekian problem of knowledge in conjunction with the institutional theory of the firm: how the knowledge within the firm grows and how it is distributed among the participants in the enterprise. An entire array of other Austrian works of the firm explores various aspects of this relationship between knowledge and organizational structures (Foss, 1999; Tsoukas 1996; Minkler, 1993).

The problem that I want to analyze in this paper is much narrower: it concerns some limited (in scope) refinements of the orthodox Austrian theory of the firm – the one combining Coase's concept of the firm with the Austrian arguments about entrepreneurship and prices. This new approach most notably pursued by Nikolai Foss and Peter Klein (Foss et al, 2007; Foss and Klein, 2012) strives to include into the analysis the theory of capital and to show that initial synthesis of Coase and Mises was justified by arguing that the Coasean notion of the firm is the best environment to accommodate the Austrian theory of heterogeneous capital, developed by Menger and Bohm-Bawerk. This is a broadening and strengthening of the theory insofar as it implies that there is a deeper structural affinity between Coaseanism and Misesianism, which concerns the very heart of the Austrian theory – the

theory of capital. As we know, the main dividing line between Bohm-Bawerkian and neo-classical theories of capital was over the issue whether capital should be conceptualized as a homogenous fund of value, consisting of identical and perfectly substitutable elements or rather as a collection of heterogeneous, imperfectly substitutable goods that incur significant costs of changing the employment of different parts of the capital stock over time. The latter position is accepted by Foss and Klein, and it is argued that the firm understood as a Coasean planning-hierarchical organization is needed in order to manage the costs of re-utilizing the heterogeneous capital goods. To further broaden their theory, Foss and Klein argue that the capital owner's managing of capital goods is necessary in order to discover the new attributes of those goods amenable to different prospective entrepreneurial utilizations (Fos et al, 2007). In a sense, Klein and Foss theory attempts to combine Misesian and Hayekian-Kirznerian concepts of entrepreneurship and to merge that composite with the Coase's theory of the firm.

I criticize this project in two steps. First, I offer an analysis of the reasons why the general attempts to synthesize Coase's firm and Mises' entrepreneur are untenable. In the second part, I follow up with an analysis of why and how Foss-Klein refinement actually adds just another layer of mistakes to the originally mistaken idea that Misesian entrepreneur needs the firm understood as a planning hierarchy.

2. COASE'S THEORY OF THE FIRM AND MISESIAN ENTREPRENEURSHIP

One of the most critical facts about the Coase's theory of the firm, which is very seldom emphasized, is that it represents just a special application of the general market socialist and market failure theories (Bylund, 2014). Coase writes his famous paper in 1937 in the middle of the controversy over the possibility of rational economic planning in socialism.² And he sides squarely with socialists and against Mises and Hayek in this debate by accepting the perfect competitive theory as the standard for judging the efficiency of the market. Coase tries to rehabilitate central planning as a mechanism of coordinating economic activity, badly damaged by the Austrian critique in the 1920s and 1930s.

This is quite obvious at the very beginning of the article when Coase contrasts price mechanism and planning, and ponders their relative usefulness:

It is easy to see when the State takes over the direction of an industry that, in planning it, it is doing something which was previously done by the price mechanism.

2 Useful reviews of the debate are given in Rothbard (1991), "The End of Socialism and the Calculation Debate Revisited", *The Review of Austrian Economics*, Volume 5, Issue 2, pp 51-76; and Kirzner (1988), "The Economic Calculation Debate: Lessons for Austrians", *The Review of Austrian Economics*. 1988, Volume 2, Issue 1, pp 1-18.

What is usually not realised is that any business man in organising the relations between his departments is also doing something which could be organised through the price mechanism. There is therefore point in Mr. Durbin's answer to those who emphasise the problems involved in economic planning that the same problems have to be solved by business men in the competitive system...The important difference between these two cases is that economic planning is imposed on industry while firms arise voluntarily because they represent a more efficient method of organising production, In a competitive system, there is an "optimum" amount of planning!" (Coase, 1937: 3)

Thus, Coase argues that an efficient economic system represents some kind of market socialist combination of planning and price mechanism; what the entrepreneurs do is the same thing that the central planners in Russia are doing. The only difference is that socialists may have pushed things too far, whereas firms in capitalism provide for an 'optimal' amount of planning.

To what extent is Coase reiterating the position of market socialists, primarily Lange and Lerner in the debate about the rational planning in socialism, is best seen by the fact that he identifies the model of perfect competition as a relevant description of the process of market exchange, any departure from which is seen as a market failure. He then identifies a problem with one of the conditions of perfect competition, namely the absence of free prices on the real market (as opposed to the model) and then argues that this requires a limited central planning through the firms to remedy this market failure. This is the essence of the famous Coasean formula about the firm as a tool of "minimizing transaction costs": "The main reason why it is profitable to establish a firm would seem to be that there is a cost of using the price mechanism. The most obvious cost of "organising" production through the price mechanism is that of discovering what the relevant prices are. This cost may be reduced but it will not be eliminated by the emergence of specialists who will sell this information." (Coase, 1937: 4).

Coase clearly believes that the only imaginable situation in which the price mechanism would operate efficiently is the one conforming to the preconditions of perfect competitive equilibrium, including free information. Since this condition is not met in reality, the price mechanism 'fails', and the firm represents a correction of this 'failure'; planning has to come to the rescue by its intervention in the form of the entrepreneurial firm, which performs exactly the same kinds of function all other government interventions are designed to perform in all market failure theories: bringing about the conditions of perfect competitive markets. The real markets are deficient, because of the transaction costs of using the price mechanism and "planning" corrects this deficiency.

Coase sees the price mechanism and "planning" as two competing coordinating mechanisms that peacefully coexist. Depending on the specific circumstances of a given society, one of them could have been dominant or 'supersede' the other in a certain area or certain branch of industry, but from his point of view there is no necessary contradiction between the two. As he emphasizes, in an economy there is always equilibrium between the marginal

utility of an additional transaction carried out within a firm and marginal utility of an additional market transaction (Coase, *ibid.* p. 405). Moreover, which one of them would relatively prevail primarily depends on the technical characteristics of the industrial structure in question. Some technological innovations could drastically increase the marginal gains of using entrepreneurial, 'non-market' solution, while the costs of using the market are stagnating: "Inventions which tend to bring factors of production nearer together, by lessening spatial distribution, tend to increase the size of the firm. Changes like the telephone and the telegraph which tend to reduce the cost of organising spatially will tend to increase the size of the firm. All changes which improve managerial technique will tend to increase the size of the firm." (Coase, *ibid.* p.397).

It follows from here that market is understood almost like a nuisance or a necessary evil, which characterizes the less technologically advanced industrial stages: the more the technological and organizational capabilities of a society increase, the more the firm becomes able to 'supersede' the market and the price mechanism. Coase even offers a telling example of Russian emancipation from serfdom and the zigzag movement from "central organization" to "market" and all the way back again, all depending on the technical and technological characteristics of the production process: "It seems important to realise that the passage from the domestic system to the factory system is not a mere historical accident, but is conditioned by economic forces. This is shown by the fact that it is possible to move from the factory system to the domestic system, as in the Russian example, as well as vice versa. It is the essence of serfdom that the price mechanism is not allowed to operate. Therefore, there has to be direction from some organiser. When, however, serfdom passed, the price mechanism was allowed to operate. It was not until machinery drew workers into one locality that it paid to supersede the price mechanism and the firm again emerged" (Coase, *ibid.* 397).

Therefore, the capitalist firm is just a new form of the same model of organizing production known from slavery or serfdom, the form which re-emerged as soon as the favorable economic and technological conditions allowed for the formation of the factory system. An untrammelled "market" dominated by the price mechanism was just a short and relatively insignificant intermezzo.

Having all this in mind, it is very peculiar that the mainstream of the Austrian school accepted wholeheartedly Coasean theory of the firm and moreover attempted to integrate it with the Austrian models of monetary calculation, capital heterogeneity, and economic entrepreneurship. This starts with Mises and Rothbard who were first to appropriate the Coasean notion of the firm as a hierarchy and a planning instrument, and then the same error was echoed and further amplified by the whole host of their followers (Klein, 1992; Foss, 1996; Ioannides, 2002; Langlois, 1995; Klein and Foss, 2012). The basic and obvious problem with all the attempts to reconcile the Austrian theory of monetary calculation with Coase's theory of the firm is that the latter had been devised as one of the arguments against the former. And the same thing that Coase explains as beneficial and remedial features of the commanding authority within the firm is described by the Austrians as a form of economic inefficiency and chaos; after all, what else "superseding the price mecha-

nism” by planning could be within the Austrian framework? For example, Coase this way describes the operation of the firm:

As D. H. Robertson points out, we find islands of conscious power in this ocean of unconscious co-operation like lumps of butter coagulating in a pail of buttermilk.” But in view of the fact that it is usually argued that co-ordination will be done by the price mechanism, why is such organisation necessary? Why are there these “islands of conscious power”? Outside the firm, price movements direct production, which is co-ordinated through a series of exchange transactions on the market. Within a firm, these market transactions are eliminated and in place of the complicated market structure with exchange transactions is substituted the entrepreneur-co-ordinator, who directs production. (Coase, 1937: 388).

Compare this now with how Murray Rothbard explains economic irrationality of socialism, with price controls and government ownership:

...islands of non-calculable chaos swell to the proportions of masses and continents. As the area of incalculability increases, the degrees of irrationality, misallocation, loss, impoverishment, etc., become greater. For each governmental firm introduces its own island of chaos into the economy: there is no need to wait for full socialism for chaos to begin to work. (Rothbard, 2009: 953)

We here clearly see that Rothbard and Coase describe essentially the same phenomenon, the spread of the non-market, planning coordination within a free market environment, using the same imagery of “islands” and seas. But, the differences in assessment of this phenomenon could not be more striking: Coase sees the ‘islands of planning’ as a welcome corrective to the inefficiencies of the market system reflected in high transaction costs, whereas Rothbard, on the contrary, sees in the same islands of planning the areas of ‘incalculability’ and economic irrationality, which, if allowed to spread over a sufficiently large part of the economy, could wreak havoc on it, even before we get to the point of full socialism (which is just a state in which there is nothing but irrationality of planning, encompassing continents now, instead of mere ‘islands’).

It is nothing short of astonishing that irrespective of this, Rothbard accepts Coase’s theory of the firm! And, moreover, he offers the Austrian arguments on price calculation as a supplement to Coase’s theory in determining the optimal size of the firm.³ Namely, according to Rothbard, the firm as a central planning island in the sea of price coordination could exist only insofar as the independent, outside markets for capital goods are preserved that would protect the integrity of the internal pricing mechanism within the firm (Rothbard, 2009).

3 See Rothbard, 2009, pp. 609–617, and further development by Klein 1996.

The most peculiar contradiction here is that Rothbard in the same breath accepts Mises' argument about the impossibility of rational economic planning absent the market prices, and Coase's theory about central planning as a way to correct the failures of market prices! But, he cannot have it both ways: non-market "planning" is either a problem, or a solution; it cannot be both: either it is an 'island of incalculability' which, if it grows enough, destroys the economic system (but even at lower levels of intensity, it impairs the economic efficiency) or it is a transaction cost-minimizing tool that helps overcome the imperfections of the market mechanism. There's no way of squaring the circle here.⁴

One could say that Rothbard diverges from Coase insofar as he identifies the calculational chaos with an increase in government control and ownership over the resources, whereas Coase did not make such a distinction. Not the firms in general, but government-owned firms are the ones that increase the calculational chaos according to Rothbard. But, this is irrelevant. From the Coasean point of view, it does not matter at all whether we are talking about the government-owned or private firms; both are equally the instruments of organizational planning that equally replace and supersede the market. Rothbard never tried to refute this; so, he is caught in a vice: if he wants to draw a sharp distinction between the government-owned firm and the private firm in the context of economic calculation, then he has to abandon the notion of the private firm as a planning entity (which he does not want to do); or if he wants to retain the Coasean notion of any firm as an organizational entity 'superseding' the price mechanism, then he has to concede that calculational chaos equally increases with the emergence and growth of both the private and government-owned firms. Then the only consistent solution is to look at every single firm, private or not, as an impediment to market efficiency.⁵ He cannot constantly equivocate by defining "the firm" as a "private firm" when trying to demonstrate the supposed advantages of this form of organization, and as "government firm" when trying to demonstrate the 'calculational chaos' to which any central planning inexorably leads.

4 The only attempt to do this, as far I can tell, is Lewin (1998). But, although the paper promises to answer the following questions: "On the one hand, if socialism is indeed irrational, in the sense of precluding the ability to perform the necessary calculations, how is it that the firm is not similarly encumbered? After all, is not a state socialist system simply one large firm? And are firms not islands of socialism in a market sea?" (Levin, 1998: 500-501). The ensuing analysis does not tell us anything about this set of issues. Instead, Levin argues that the entity called the "firm" is necessary to provide an environment for entrepreneurial calculations of cost and profit. But, this is perfectly consistent with the nexus-of-contracts view of the firm and does not explain at all why we would have to understand the firm as a commanding entity, rather than as a legal fiction for voluntary contracts. Moreover, he equivocates about the very meaning of the key term "planning". At the beginning, this term means the relationship between the owner and employees. Later on, he switches to the notion of planning as entrepreneurial anticipative calculation of costs and benefits. This is close to what Mises would call the 'praxeological' meaning of planning: simply, every market agent plans, i.e. anticipates his own costs and benefits. If this is so, it's not clear why then the employees would not be "planning" as well. And if everybody is "planning", what is then the point of making an analytical distinction between the "planning" and the "market" in the first place?

5 This is the criticism that Demsetz (1988; 2011) levels against Coase.

3. HETEROGENEITY OF CAPITAL AND THE FIRM

We now turn to the refinements of this Rothbard's mistaken acceptance of the Coasean theory of the firm by the authors who tried to incorporate the Austrian discussion of the capital heterogeneity into the system (Klein and Foss, 2012, Foss etl 2006;). It is safe to say that what they achieved was simply to generalize and amplify Rothbard's initial mistake of treating central planning as a cure for market failure by infecting even the Austrian capital theory with Coaseanism.

Let's start by discussing the differences between the Austrian and neoclassical capital theories, and then explore what conclusions the Rothbardians derive from those differences in the context of the theory of the firm. Heterogeneity of capital stems from the general analysis of value, price, and the market process by Menger and Bohm Bawerk. Carl Menger revolutionized economic theory by reinterpreting the origins of value and reversing the flow of economic causation in his famous book 'The 'Principles of Economics', as compared to earlier theories. Unlike his classical predecessors, Menger argued that the source of value is individual human valuation of consumer goods' utility for the satisfaction of his needs, and that the standard of evaluation is marginal utility which is decreasing (Menger, 2007). This helped solve the paradox of value that plagued classical economics, mired in the cost of production and labour theories of value, which were unable to explain the market value of non-reproducible goods, or natural goods.

However, the critical corollary of this revolution in the theory of value was Menger's theory of imputation that shifted the causal arrow from the classical cost theories, by claiming that the cause of all material production and the source of their value are human needs and human evaluations of the material goods, and that all other goods that are used at some point to produce those final consumption articles derive their utility and their price from the value of the final goods. Marginal utility determines the price of the final goods, whereas the marginal productivity in producing those final goods determines the price of intermediate capital goods. The value is, so to speak, imputed backwards to the capital goods and land from the consumer goods, rather than prices of final goods being determined by the objective, ex ante given costs of their production (Menger, 2007: 51-74).

A further consequence of this shift is that production is not seen any more through the lenses of an input-output analysis, depicting the process of production as a transformation of a given set of inputs (factors) into the set of final goods, but rather a complex intertemporal process that starts with the higher order goods, as Menger names them, and then proceeds through a series of intermediate steps to the final stage of consumption goods (Menger,ibid.).

Bohm-Bawerk describes this as a "roundabout" process of capitalist production, and at each level of production a large-scale process of selection and specialization of capital goods is taking place; the further we go from the stage of final goods along the chain of production structure, the less specialized goods become: land, raw materials or products of nature could be put in very many uses; some intermediate goods such as the refined metals and some machines are less specialized than some others such as more narrowly special-

ized capital equipment or half-products (Bohm-Bawerk, 1959: 3–15). And this is the source of heterogeneity of capital in the Austrian theory.

The most obvious feature of this theory puts it at odds with the now conventional neo-classical understanding of capital. All those theories, in various forms depict capital as a homogenous fund of value which serves the purpose of automatically converting the inputs into the output. The production process is understood as instantaneous and capital goods as perfectly substitutable. The factors of time and heterogeneity of inputs are done away with.⁶

4. KLEIN AND FOSS REINTERPRETATION

One aspect of the Austrian capital theory we are here most interested in is how it interacts with the conventional and nonconventional theories of entrepreneurship and the firm. In the Austrian literature, this issue is pursued with vigilance in the last couple of decades, and we are going now to explore some of the results that are achieved and their limitations. As already suggested, the Austrian literature pursuing those kinds of complications is mostly Misesian-Rothbardian in its direction, and the exposition and critique here would be directed to the entire tradition, although I would single out a couple of the most prominent and canonical authors of this tradition. Primarily, I have in mind the refinements and broadening of the paradigm brought about by Nicolai Foss and Peter Klein, and some other authors following their lead.

The basic thrust of their argument is an attempt to broaden the application of the Rothbardian dubious reconciliation of the Coasean “organizational” theory and Austrian price theory by including the theory of capital into the synthesis. And although impressive in its learning and comprehensiveness, this attempt is untenable for the same reasons Rothbardian initial argument for Coaseanism is wanting. Klein and Foss (2012) as well as Foss et al (2007) start from the fact of heterogeneity of capital in the Austrian tradition and contrast it with the treatment of capital in the mainstream neoclassical tradition. This is, in a sense, an already explored possibility; for example Dupleco and Garrouste (1999) are trying to bridge the gap between the Hayekian-Kirznerian entrepreneurial knowledge and the Bohm-Bawerkian asset specificity. However, they use the firm just as a functional black box to perform the feat, avoiding any explicit Coasean themes for analysis of the inter-firm structure. They are not interested in the structure of the firm.

6 For a good summary of the debates on capital theory between Austrians and neoclassicals see Hayek (1935) and Cohen (2008). The debate was waged over the period of more than 30 years. For the Austrian expositions see, Bohm Bawerk (1895), Hayek (1935) and Machlup (1935). For the neoclassical “fund of value” position see Clark (1895) and Knight (1934).

Klein and Foss, on the contrary, make the Coasean structural theory, appropriated by Mises and Rothbard, the center piece of their synthesis, and then build further from there by enriching the synthesis with the Kirznerian and Bohm-Bawerkian themes. Their approach is to correct the neoclassical theory not only in one point (the structure of capital) but in two: the structure of capital and the structure of the firm. First, the functional theory of the firm which serves to transform the given set of inputs into outputs is simplistic and actually wrong since it depicts the process of production as a one-stage enterprise; capital is instantaneously transformed into a set of final, consumer goods. Foss et al argue, in line with Bohm Bawerk that production process unfolds in the numerous stages, and is diachronical rather than instantaneous. Also, in order to understand the entire process of production, it is not sufficient to postulate the firm as the production function, but one needs to explain its internal structure and working: when does firm emerge, how and why it grows, when it ceases growing, why certain transactions are carried out within the firm and certain others outside of it, and so on.

Foss et al (2007) attempt to establish the rationale for Coasean firm as the only solution for those problems by using the tools provided by Frank Knight. Namely, they emphasize the importance of entrepreneurial judgement which is not based on any given criterion or probability distribution among the given entrepreneurial choices. Judgement is a qualitatively different kind of decision-making than 'optimizing' in the neoclassical framework; starting from Knightian distinction between risk and uncertainty, they argue that entrepreneurial decision-making includes the handling of the situations which are not subject to ordinary risk, and hence uninsurable. An entrepreneur deals with what Knight calls uncertainty and Mises the class probability (as opposed to case probability). This is a set of circumstances in which the array of possibilities is not closed and a great deal of what the entrepreneur does has to do with the anticipation of future gains from qualitatively different possible entrepreneurial initiatives with unpredictable results: "Judgment primarily refers to the process of businessmen forming estimates of future events in situations in which the relevant probability distributions are themselves unknown. Entrepreneurship represents judgment that cannot be assessed in terms of its marginal product and which cannot, accordingly, be paid a wage." (Foss et al, 2007: 1167).

The second great innovation that Foss and Klein offer is the claim that in the world of homogenous capital goods of neoclassical theory, the identical and perfectly substitutable "shmoos", in the parlance of Paul Samuelson, there would be no need for corporate governance, industrial organizations such as firms or any other non-market (in the Coasean sense) forms of governance: "In a world of shmoo capital economic organization is relatively unimportant. All capital assets possess the same attributes, and thus the costs of inspecting, measuring, and monitoring the attributes of productive assets is trivial. Exchange markets for capital assets would be virtually devoid of transaction costs" (Foss et al, 2007: 1168). Hence, the very existence of the firm is dictated not only by standard Coasean transaction costs but also by Bohm Bawerkian capital transaction costs.

However, Foss and Klein refine the theory further yet by making it more subjectivist and less 'Ricardian'. Heterogeneity of capital is not understood only in the standard Menger-

Bohm Bawerkian way, as a reflection of the multitudes of purposes to which any physical piece of equipment or factor could be put, but also includes the concept of heterogeneity of a single and seemingly homogenous capital asset. Using Kirzner's and Hayek's conceptualizations of the process of market competition, they argue that capital goods are not just separate and purpose-specific physical things whose replacement or reutilization is costly, and hence it does not make much sense to talk about capital as a homogeneous fund of value; even more importantly, a single physical asset is a bundle of numerous attributes, many of which are not always understood or utilized as economic goods (Foss and Klein, 2012: 116-131). So, heterogeneity of capital goods is also seen in the fact that different entrepreneurs have different ideas of how a multitude of attributes that any physical item has could be combined and exploited to make profit. Utilization of capital goods is a continuous and never-ending process of discovery of the useful characteristics of every single capital asset that is used in the production process. The same capital good could be radically different to the two different entrepreneurs, depending upon which concrete attributes any one of them perceives as important for the production processes at hand. A very large number of entrepreneurial discoveries consist not in finding a new material or a new capital good to serve the given purposes (or new purposes for the same attributes of a given good), but in discovering some new feature of the already existing capital asset which had not been perceived before as suitable for such purposes.

This way, Klein and Foss combine in a sense the best of both worlds; from the Misesian, essentially Bohm Bawerkian, tradition they take the concept of entrepreneurship as economic utilization of heterogeneous physical capital, as human action on the real markets which is concentrated on production, exchange, investment using monetary calculation and market prices to allocate physical resources to alternative uses. From the Hayekian and Kirznerian theory of competition as a discovery procedure, and the entrepreneur as a knowledge economizer, they take a radical subjectivist emphasis on specific tacit knowledge and gradual discovery of unknown data to develop the notion of capital goods as being subjectively structured by the information, knowledge, and perception of the entrepreneurs themselves. That way, heterogeneity of capital is a pervasive fact of economic life; it is not seen only in objectively different physical characteristics and purposes to which different goods could be put, but also in subjective attributes that different entrepreneurs assign to the physically same capital goods.

The function of the entrepreneur in this picture is to embark upon a process of production with physical capital by using his own judgement in the conditions of radical uncertainty in order to discover a set of attributes of the given assets, allowing him to provide the marginal value to the consumer. Their emphasis on capital theory and their ingenious reinterpretation of the concept of heterogeneity allows Klein and Foss to offer a brilliant synthesis of the Misesian and Hayekian perspectives within the theory of entrepreneurship. On the one hand, an entrepreneur is not just the discoverer and arbitrageur who passively reacts to the changes in market data and by his prompt and skillful use of his specific and tacit knowledge beats the competition in the race for profit. An entrepreneur is also someone who directs and employs physical capital, the transformation of which takes time, cop-

ing with uncertainty and requires judgement and planning. Organizing physical production is an even more important element of entrepreneurial activity than a mere epistemological function of price and opportunity discovery. The conceptualization of the unity of those two aspects of entrepreneurial activity is a signal contribution of Klein and Foss.

5. CRITIQUE

The gist of the Foss/Klein revision of the standard Austrian framework is best expressed in their previously quoted statement “In a world of shmoo capital economic organization is relatively unimportant. All capital assets possess the same attributes, and thus the costs of inspecting, measuring and monitoring the attributes of productive assets is trivial. Exchange markets for capital assets would be virtually devoid of transaction costs” (Foss et al, 2007: 1168). This statement is essentially just a translation of the basic Coasean claim about the absence of the firm in the neoclassical paradigm into the language of capital theory. Coase imagines that the only reason why the firm could emerge is to minimize the transaction costs,⁷ and since the transaction costs in the neoclassical model are zero, then neoclassical theory ‘fails’ to explain the very existence of the firm. In the same manner, Foss et al claim that the only reason why the “organizational forms” of governance could emerge is to deal with the transaction costs of operating the capital assets, and since the capital assets are identical in the neoclassical treatment, the transaction costs are minimal and hence the firm is unnecessary.

But, both claims are equally dubious. As Harold Demsetz (2011a) demonstrates, the firm in the neoclassical model is meant to perform a functional role as a part of the system of exchange between the open market and households. Firms emerge not because the market is characterized by high transaction costs but because it is characterized by a high extent of division of labor; the reason for a firm’s existence is that it provides a large degree of specialization which increases productive efficiency and represents an alternative to the self-sufficient provision within a household. The firm is an instrument of production and exchange which becomes more important as the extent of market specialization and division of labour increases, not less. The opposition in the neoclassical theory is not the one between the market driven by price signals and the firm driven by commands and planning, but the one between the markets with the firms as the instruments of specialization and exchange and households as self-sufficient and non-market environments of production and distribution.

Contrary to what Coase claims, the firm could exist even in the zero transaction cost world, and moreover it is likely that the extent and the reach of the firm production in such a world would have increased rather than decreased, because it would have made the provi-

7 For a critique of this assumption see Demsetz 1988.

sion of goods within the household less appealing (since the cheaper firm-produced goods are more widely available than otherwise), and would have, hence, allowed the firms to serve a much wider circle of people (Demsetz, 2011a: 10). Coase completely neglects the productive efficiency in studying the origins, boundaries and growth of the firm. The only costs allowed in his model are transaction and managing costs, and the balance of those two costs on the margin determines the extent of the firm's growth and size. But, we can have the enormously large firms in the world of zero transaction costs because of a higher degree of specialization and productive efficiency which makes the production for the market more profitable than self-sufficient production. A firm can continue growing if the transaction costs are zero just as it can shrink if the transaction costs skyrocket.

Foss et al (2007) make the analogous mistake in their treatment of the 'shmoo' capital. Just as Coase wrongly imagined a world of zero transaction costs in which all the market agents would just trade and exchange goods using the price mechanism, Foss et al imagine a world of perfectly substitutable capital assets in which the entrepreneurs would cease to control and own any capital goods and would communicate simply through the spot contracts on the market, making firms all but disappear: "The possibility of specifying all possible uses of an asset significantly reduces the costs of writing complete, contingent contracts between resource owners governing the uses of the relevant assets. Contracts would largely substitute for ownership, leaving the boundary of the firm indeterminate" (ibid. 1168).

But, this also completely neglects the effects of specialization and productive efficiency: in the world of perfectly identical and substitutable capital goods, the production within a firm would become infinitely more rather than less appealing. If we accept that the firm allows for greater specialization and a decrease in per unit costs, then the world of shmoo would be an Eldorado for the firm as an organization. Heterogeneity of capital is a cost that the intra-firm production faces, rather than an incentive to it. It makes the functioning of the market system less smooth and more complicated than it would have otherwise been; if heterogeneity is assumed away, i.e. the assets are assumed not to be specific and substitution easy and quick; then the firm could exploit all the advantages of production that it already has, but on a much grander scale. In the world of shmoo the firm would be much more important and widespread than it is now. Transaction costs, however we define the term, represent just one element of the environment in which the firms operate, whereas productive efficiency, the extent of the market specialization, oftentimes represents more important factors influencing the size and growth of the firm. In other words, in the world of shmoo capital, the transaction costs of switching the uses of a given capital good would be negligible, which would provide a tremendous incentive to the firms to widen the scope of their operation, instead of shrinking it. Lower transaction costs of exchanging the capital assets across the stages and branches of production would facilitate the cheap production within the firm, reducing the unit costs and increasing the productive efficiency. Just as in the case of Coase's theory, households would minimize the extent of self-sufficient production and increase their reliance on the outside markets, i.e. on buying the cheaper products from the firms.

It has already been observed (Demsetz, 1988; Hulsmann, 2004; Jankovic, 2010) that the very concept of transaction cost lacks any serious explanatory and predictive power.

Limited knowledge, costs of monitoring, difficulty in discovering prices – all those diverse forms of costs are arbitrarily subsumed under the generic rubric of ‘transaction costs’. But, all these cognitive and behavioral limitations of human beings are with us with or without the firms, and it is not clear why only those costs would be relevant and not some others that Coase did not specify. Is lifting a pen to sign a contract transaction cost or not? Or waiting for the elevator that would take us to the floor where the firm headquarters are? (Hulsmann, 2004) Everything and anything could qualify as a transaction cost depending on how far we are willing to go in calling every obstacle we encounter a transaction cost. We essentially end up with a notion of ‘transaction costs’ which approximately means: ‘any kind of reason or problem that we can think of why a particular transaction had not taken place’. ‘Transaction costs’ become indistinguishable from the ‘costs of resolving problems’ (Demsetz, 1988:) As Demsetz bluntly puts it: “A cost is cost whether we name it transaction cost or fertilizer cost. Consider a manufacturer who does not ship his product to another firm because the other firm is unwilling to pay the costs of producing the product. What difference does it make to efficiency if the high cost is due to the price of labor or to the cost of shipping the product – and is not the cost of shipping the product a transaction cost?” (Demsetz, 2011b: 13). In the final instance, this entire line of reasoning leads to the claim that we need hierarchical planning and commands to replace market prices since we live in an imperfect world.

Foss et al just add one more item to this unending list of poorly defined “transaction costs” – the costs of shifting various specific, heterogeneous capital goods from one use to another. They assume the same Coasean critique of the markets in which the precondition of its efficient functioning is the absence of transaction costs. In this context, the requirement is that capital be the homogenous blob, shmoo capital, seamlessly fluctuating between the various stages of production and various branches of industry: management, entrepreneurship, vertical integration of production, all those phenomena the Austrian Coaseans see as the effect of capital being heterogeneous and specific. Somewhat crudely, but correctly put, they see, like Coase, central planning as a mechanism for coping with imperfect information and limited knowledge.

However, although this new theory faces the same kind of critique we developed for the original Coasean position, the situation is much worse for Klein, Foss and others than for Coase. Namely, Coase never claimed that his purpose was to vindicate Misesian theory of economic calculation and entrepreneurship; on the contrary, as we have seen, his paper could be read at one level as a polemic against Mises and Hayek. Austrian Coaseans on the other hand tied themselves in a knot by claiming on the one hand that market prices are generally superior to central planning, and at the same time that central planning is superior to the market prices within the firm. However, they follow Rothbard in the failure to provide any Austrian justification for the notion that planning and non-market commandeering could be more efficient than a market in any circumstances. You can claim that Coase is right in assuming that transaction costs represent the market failure that has to be remedied by planning, or you can claim that prices are always better than non-market commandeering. But, you cannot claim both in the same time!

Contrary to Hayek and in accordance with Coase, they see the perfect information and zero transaction costs as the conditions of market price efficiency; only in the world of sh-moo economic coordination would be carried out by contracts and market transactions. In the real world of positive transaction costs, we need 'planning'. But, is not the basic Hayek lesson that we need price system and market competition exactly because we do not live in a world of perfect information, and capital goods are not sh-moo? The fact that information in reality, as opposed to the assumptions of the model, is imperfect and limited, prompted many economists to believe that market coordination was inefficient and had to be supplemented or supplanted by the conscious and deliberate control of economic activity in order to approximate the economic outcomes to the properties of the model. Hayek's critique of that conception was that the very essence of market is to economize and optimize the use of scarce resources including knowledge: free competition is the best mechanism so far discovered allowing individual people with limited information and bounded rationality to extract from society more knowledge about the appropriate use of resources than any one of them could have possibly possessed in isolation (Hayek, 1945). The very essence of free competition is to discover the new knowledge about the ways to use and combine the resources, not in applying the knowledge that we already have. The less perfect the information is, the more we need the market prices and competition! In a world of perfect information, we would not need markets and prices at all. We have them precisely because information is costly, scarce, and limited.

By the same token, the fact that capital goods are heterogeneous does not mean at all that the mechanisms of control and transformation of those goods into the final consumer products should be guided to a larger degree by a conscious authority and withheld from the domain of market transactions. The fact that there are positive transaction costs of changing the modes of utilization of capital goods (the costs of 'asset specificity' or of 'attribute discovery') does not mean that the market in capital goods is 'inefficient' or that we need to supplement it by some form of central planning to correct its failure. The fact of specificity-heterogeneity of capital goods is equally irrelevant for justifying the concept of a firm as hierarchical organization as the imperfection or asymmetry of information is irrelevant for justifying government regulation of the insurance business. The more heterogeneous the capital goods are, and higher the transaction costs of switching the uses of those goods are, the "thicker" the market in which the entrepreneur operates has to be. He needs more information, not less, to deal with the more complicated situations, and the firm understood as a central planning instrument would, by definition, provide less information than the 'market'. This does not mean that in the world of more heterogeneous capital we would need firms less, but that firms are not central planning entities.

As we have seen, Klein and Foss enrich the theory tremendously by combining in very creative ways Kirznerian, Knightian and Bohm-Bawerkian themes with the old Rothbardian and Misesian theories. However, the final result is, paradoxically, to even more undercut the initial Rothbardian position, instead of strengthening it. For example, if heterogeneity of capital is not just the old, good, run-of-the-mill "asset specificity", but also includes the subjective evaluations of the various attributes of the same goods, then it is not clear what pur-

pose the notion of transaction costs as a cost of switching the uses of single heterogeneous goods could play in the entire scheme. If transaction costs mean the costs of unpredictable discovery of the new features of a *given* good, what is the basis for belief that this discovery would be more likely to occur within the Coasean firm rather than outside it? One million entrepreneurs cooperating via prices and spot contract could do an equally good job of discovering the new features of capital goods as can do the entrepreneurs organizing the team production. There is no basis, at least in the framework developed by Foss and Klein, to believe otherwise. Instead of offering further support for Austrian Coaseanism, Foss' and Klein's refinement actually additionally illustrates its unviability.

Another downside of this refinement is that it makes some problems vis-à-vis ownership over the firm that the Austrean-Coasean theory has had even worse. Namely, in the standard Austrian theory a very sharp distinction is made between the entrepreneur and manager; The entrepreneur is a creative force which initiates business plans, takes up and recovers credit, makes decisions about what production processes to initiate, how much money to invest here or there, how many workers to employ, whether to expand or scale back the operations of the firm and so on (Mises, 1998). A manager, on the contrary, is limited to the delegated and routine tasks. In addition to this, the entrepreneurial function is tightly connected to the ownership of the firm: "the entrepreneur is nearly always also a capitalist and the capitalist is also an entrepreneur", says Peter Klein (2010). This means that entrepreneur must hold ownership over the assets whose different attributes are to be discovered. Klein and Foss (2012: 120) emphasize "ownership is a low cost form of allocating the rights to attributes of assets that are created or discovered by the entrepreneur-owner." In Foss et al (2007) they use a similar argument, but expressed in Knightian terms: "...there is no market for the judgment that entrepreneurs rely on, and therefore exercising judgment requires the person with judgment to start a firm...Judgment thus implies asset ownership, for judgmental decision-making is ultimately decision-making about the employment of resources. An entrepreneur without capital goods is, in Knight's sense, no entrepreneur." (Foss et al, 2007: 1167)

One obvious problem with the identification of owner and entrepreneur is that all the essential functions of the alleged entrepreneur-owner in the modern corporation appear to be performed by paid employees, rather than by the owners themselves. In the modern corporation of a great mass of individual shareholders very few of them had any idea about the production process or the true business of the firm, and they instead hire the other people to lead the corporation and organize the production process at all stages. In most cases, those owners are not only excluded from making and executing the business plans, but have actually a limited liability for the corporate performance. This is really awkward: if it's true that entrepreneur=owner, then the corporation represents a really anomalous form of organization which should be expected to have a dismal record: what kind of organization plagued by such moral hazard in which the owner reaps the benefit of a successful business while being shielded from at least some of the consequences of failure could be efficient? And yet, the modern corporation seems to dominate the business world. Either it is somehow artificially privileged by the legislation, or the concept of the entrepreneur

being identical to the owner does not hold, i.e. it must be that “managers” could exercise successfully the entrepreneurial function.

Klein and Foss try to solve this obvious problem by claiming that there are two different types of entrepreneurial judgement: the “original” judgement exercised by the owners of a firm and “derived” judgement exercised by the managers and other people to whom the rights to decide about certain issues are delegated by the owners (Klein and Foss, 2012). Obviously struggling with the fact that the CEOs and other senior managers of modern corporations make the most important decisions about the business operations of the firm, they redefine the entrepreneur: this is not anymore a person who owns the corporation, makes the business plans as well as all other critical decisions (as in earlier Austrian literature), but only the person or persons who have a residual or ultimate decision making power. In other words, he can fire the CEOs or stop any particular business plan.

This solution may or may not sound convincing in and of itself, but one thing is certain: it is very difficult to square with the “attributes discovery” theory of entrepreneurship. For the purposes of making the dubious Austrian identification of entrepreneur and owner look more realistic, Foss and Klein are forced to empty the notion of entrepreneur of almost any concrete content, apart from the ultimate control of assets, which is little more than an analytic definition of the term “owner”: their “clarification” that the owner is entrepreneur because he controls the assets amounts to little more than a claim that owner is owner because he is the owner. And yet, while developing their theory of ownership as a “low cost form of allocating the rights to attributes of assets that are created or discovered by the entrepreneur-owner” they seem to have in mind a much richer notion of the entrepreneur, who actually *personally does* the job of attribute discovery and is rewarded for this by the ownership over the attributes he discovered. In other words, in the theory of attributes discovery, the concept of the entrepreneur is much closer to the person exercising “derived judgement”, i.e. the manager, rather than to the ‘real’ entrepreneur with his “original judgement”. So, the entire theory is in a sense based on an equivocation: when the problem of corporation should be solved, the concept of entrepreneurship is emptied of most of its usual content and identified with a minimalistic “original judgement”; on the other hand, when the theory of “attributes discovery” is developed this standard content is smuggled back again into the model.

The basic weakness of the theory mimics the general weakness of the old Austrian-Coaseanism of Mises and Rothbard: the impossibility to square the logic of individual entrepreneurship, market prices and market process, with the socialist concepts of organization, hierarchy and commandeering taken from Coase. In Foss’ and Klein’s revision, this is even more evident since they emphasize the concept of entrepreneurship, as well as the heterogeneity of capital, which both fit very nicely the general Misesian paradigm, thereby making the artificial and convoluted nature of the Coasean firm even clearer. Foss’ and Klein’ theory, if I am allowed such a metaphor, resembles a man wearing a tailcoat, a cylinder, the black pants, a white shirt with a bow tie, and -muddy military boots, instead of shoes.

6. CONCLUSION

The basic claim of this paper is that the only way of making sense of the notion that a free market is superior to central planning in the context of a world characterized by heterogeneity of capital is to reject Coasean theory of the firm as a centrally planned 'organization'. Contrary to some new attempts to show that heterogeneity of capital stock (asset specificity) entails various forms of opportunistic behaviour that invite the substitution of organizational forms of governance for market transactions (Klein and Foss, 2012; Foss et al, 2007), I affirm that heterogeneity is actually an additional reason why the market prices are indispensable and why organizational and planning thinking is a wrong way to conceptualize the problem. Austrian economists who tried to fuse Misesian arguments on economic calculation and market prices with the Coasean theory of the firm never addressed the basis and fundamental disconnect between these two theories: the fact that Coase argues that central planning represents a remedy for the failure of market prices efficiently to coordinate the economic activity. The newest attempt by some Austrians to use the upgraded form of Bohm-Bawerkian capital theory as a new way of reconciling the two is equally untenable. All the problems and contradictions this brand of Austrian theory faces could be avoided very simply: by rejecting the Coasean notion that a firm represents an instrument of central planning, but rather a form of voluntary and contractual market cooperation.

REFERENCES

- Alchian Armen, Demsetz Harold (1972) 'Production, Information Cost and Economic Organization', *American Economic Review* 62(5): 777-95.
- Bohm-Bawerk von, Eugene (1959) *Capital and Interest: Volume II - Positive Theory of Capital*. South Holland, Illinois: Libertarian Press.
- Bohm-Bawerk, E. (1895) 'The Positive Theory of Capital and Its Critics I.' *Quarterly Journal of Economics* 9 (January): 113-31.
- Bylund Per (2014) 'Ronald Coase's "Nature of the Firm" and the Argument for Economic Planning', *Journal of the History of Economic Thought* 36(3): 305-329.
- Clark, John Bates (1893) 'The Genesis of Capital', *Yale Review* 2: 302-315.
- Clark, John Bates (1895) 'The Origin of Interest', *Quarterly Journal of Economics* 9 (April): 257-78.
- Coase, Ronald (1937) 'The Nature of the Firm', *Economica*, New Series, 4(16): 386-405.
- Cohen, A. (2008) 'The Mythology of Capital or a Static Equilibrium: The Bohm-Bawerk/Clark Controversy', *Journal of the History of Economic Thought* 30(2).
- Cowen, Tyler and David Parker (1997) *Markets in the Firm: A Market Process Approach to Management*. London: The Institute of Economic Affairs.
- Demsetz Harold (1983) 'The Structure of Ownership and the Theory of the Firm', *Journal of Law and Economics* 26(2).

- Demsetz Harold (1988) 'The Theory of the Firm Revisited', *Journal of Law, Economics, & Organization* 4(1): 141-161.
- Demsetz Harold (2011a) 'R. H. Coase and the Neoclassical Model of the Economic System', *Journal of Law and Economics* 54(4): S7-S13.
- Demsetz Harold (2011b) 'The Problem of Social Cost: What Problem? A Critique of the Reasoning of A.C. Pigou and R.H. Coase', *Review of Law and Economics* 7(1).
- Dulbecco, Philippe and Pierre Garrouste (1999) 'Towards an Austrian Theory of the Firm.' *Review of Austrian Economics* 12: 43-64.
- Kirsten Foss, Nicolai J. Foss, Peter G. Klein and Sandra K. Klein (2007) 'The Entrepreneurial Organization of Heterogeneous Capital', *Journal of Management Studies* 44(7).
- Foss, Nicolai J., and Peter G. Klein (2005) 'Entrepreneurship and the Economic theory of the Firm', In Rajshree Agarwal, Sharon A. Alvarez, and Olav Sorenson (eds.) *Handbook of Entrepreneurship: Disciplinary Perspectives*. Norwell, Mass.: Kluwer.
- Foss Nicolai and Peter G. Klein (2012) *Organizing Entrepreneurial Judgement: A New Approach to the Firm*. Oxford University Press.
- Garrouste (2002) 'Knowledge: a Challenge for the Austrian Theory of the Firm', In Nicolai J. Foss and Peter G. Klein (eds.) *Entrepreneurship and the Firm: Austrian Perspectives on Economic Organization*, Aldeshot: Edward Elgar.
- Hayek, F. A. (1936) 'Mythology of Capital', *Quarterly Journal of Economics* 50(2): 199-228.
- Hayek, F. A. (1945). 'The Use of Knowledge in Society,' *American Economic Review*, 35(4): 519-530.
- Huelsman, Guido Joerg (2004) 'The A Priori Foundations of Property Economics', *Quarterly Journal of Austrian Economics* 7(4): 41-68.
- Ionnanides, Stavros (1999) 'Towards an Austrian Perspective on the Firm', *Review of Austrian Economics* 11: 77-98.
- Jankovic, Ivan (2010) 'Firm as a Nexus of Markets', *Journal des Economistes et des Etudes Humaines* 16(1).
- Klein, Peter (1996) 'Economic Calculation and the Limits of Organization', *Review of Austrian Economics* 9(2): 3-28.
- Knight, Frank (1934) 'Professor Hayek and the Theory of Investment', *Economic Journal* 45(177): 77-94.
- Langlois, Richard N. (1995) 'Do Firms Plan?' *Constitutional Political Economy* 6: 247-61.
- Lewin, Peter (1998) 'The Firm, Money and Economic Calculation: Considering the Institutional Nexus of Market Production', *American Journal of Economics and Sociology* 57(4): 499-512.
- Mathews, Don (1998) 'Management vs the Market: An Exaggerated Distinction', *Quarterly Journal of Austrian Economics* 1(3): 41-46.
- Machlup, Fritz (1935) 'Professor Knight and the Period of Production', *Journal of Political Economy* 43(5).
- Menger, Carl (2007) *Principles of Economics*. Auburn, Alabama: Ludwig von Mises Institute.

- Minkler, Alanson P. (1993) 'The Problem With Dispersed Knowledge: Firms in Theory and Practice.' *Kyklos* 46: 569–87.
- Mises, Ludwig von (2008) *Human Action*. Scholar's edition. Auburn, Alabama: Ludwig von Mises Institute.
- Rothbard, Murray (2009) *Man, Economy and State*. Scholar's edition. Auburn, Alabama: Ludwig von Mises Institute.
- Salin, Pascal (2002) 'The Firm in a Free Society: Following Bastiat's Insights', *Journal of Libertarian Studies* 6(3).
- Tsoukas, H. (1996) 'The firm as a distributed knowledge system: A constructionist approach.' *Strategic Management Journal* 17: 11–25.

BOOK REVIEW

INVENTING FREEDOM: HOW THE ENGLISH-SPEAKING PEOPLES MADE THE MODERN WORLD BY DANIEL HANNAN, 2013. BROADSIDE BOOKS, 416 P

In the political correctness ethos of our age, one of the worst crimes an academic might commit is to make a claim of civilizational exceptionalism for a particular *nation* – i.e. a particular group of people identified primarily by ethno-linguistic ties. Fortunately, Daniel Hannan is a politician, not an academic, and thus is free to pen a book committed to such a thesis. In *Inventing Freedom*, Hannan makes a bold, intricate case for the legal and political – and therefore economic – exceptionalism of the English-speaking countries, which he refers to as the “Anglosphere.” Hannan’s book is sure to raise the hackles of multiculturalists and the diversity lobby, while appealing to the nativist sentiments of some conservatives. Yet this book is no mere polemic offered up for party ideologues, such as might emerge from the pen of a Sarah Palin or Al Franken. Hannan’s work is of a much more nuanced, scholarly bent, and brings a deep historical perspective to issues of central importance to the field of constitutional political economy.

Hannan’s claim is that the unique political constitution¹ of the Anglosphere is responsible for its broadly-acknowledged economic success and geopolitical dominance. In other words, Hannan espouses the view that the Anglosphere countries are marked primarily by their legal and political institutions, and these institutions are the most important elements in their success. This is a bold claim, but Hannan does an admirable job articulating it with a rich, detailed historical narrative of the ancient roots, historical evolution, and modern refinement of this institutional framework, which Hannan sees as coming to full flower with the US Constitution of 1787.

Critical readers who are familiar with Hannan will be suspicious of natural biases at play in this book. Hannan represents South East England in the European Parliament, the legislative body of the European Union. Hannan, an unabashed and vocal conservative, came to international prominence when his trenchant, eloquent harangue of British Prime Minister

1 Throughout this review I will use the term “Anglosphere constitution” – with a small c – to refer, not to any particular parchment, but to the unique pattern of legal and political institutions that typify the Anglosphere countries as will be described below.

Gordon Brown went viral on Youtube in 2009.² Building on the cross-Atlantic acclaim for this speech, Hannan released *The New Road to Serfdom: A Letter of Warning to America*, in 2010. Although Hannan hails from Peru – a son of British emigrant farmers – and is trilingual in English, Spanish, and French, he is proudly English in the sense of upholding what he views as England’s unique political values. Hannan is prominent amongst the Euro-skeptic movement in Britain, promoting national sovereignty and a devolution of power to the lowest possible level of government, as opposed to the pan-European social-democratic machinations of the continental Eurocrats.

Yet while Hannan is an artful conservative spokesman with a unique talent for firing up the passions of Anglo-American conservatives, he is no mere demagogue. Hannan’s background, especially his foreign upbringing, gives him a much more nuanced and informed viewpoint, which lends credibility to his views. Moreover, his knowledge of history and his reading of the relevant scholarship is deep and broad, giving his book an erudite, academic flavor. Thus even those who walk away from this book still skeptical of Anglosphere constitutional exceptionalism won’t be able to fault Hannan for not making his arguments in a clear, calm, and compelling manner.

ANGLOSPHERE EXCEPTIONALISM: THEORY, BACKGROUND AND PRACTICAL IMPLICATIONS

Hannan’s thesis in *Inventing Freedom* revolves around a political economy argument, an historical argument, and a case study writ large. I will explore each of these in turn, then connect Hannan’s work to a broader economics and political economy literature, and finally conclude with a strong recommendation of *Inventing Freedom* for students and scholars of political economy.

THE IMPORTANCE OF INSTITUTIONS

Hannan’s political economy argument focuses on the idea that a nation’s legal and political institutions – the manner in which it makes and enforces its laws and operates its government – are of paramount importance to its economic success viewed in terms of the average citizen’s standard of living. The particularly successful Anglosphere constitution, broadly conceived, is built upon the following principles:

1. Primacy of the individual: individual citizens are ultimately sovereign over themselves and their justly acquired property, and may employ (or not employ) their persons and

² <https://www.youtube.com/watch?v=94lW6Y4tBXs>

estates in any manner they wish, so long as they do not interfere with like rights of all others.

2. Representative government: as the people, individually, are ultimately sovereign, legitimacy in government arises from them and can only be delegated to the state apparatus (courts, police, etc.) by a representative process, through which the individual citizens retain ultimate control of the governing bodies this process may establish.
3. Common law: the law is seen as ancient and universal, given by nature or God. It pre-exists the state and thus the judiciary's job is not to create law, but merely to discover and apply it. The role of the legislature is also, then, a rather modest one of affirming settled legal principles and establishing rules for hard and novel cases by statute. It is *not* the legislature's job to attempt to impose any particular vision of the ideal society by forcefully rearranging people's status and property as if they were pieces on a chess board.
4. Rule of law: the law is supreme, as reflected in the ancient dictum "no man is above the law." Rule of law tames the passions and prejudices of particular men in particular times and places, and ensures that all citizens are treated equally *under the law*, i.e. that all are entitled to the same legal procedures, such as trial by jury for capital offenses – and privileges, such as habeas corpus and the "right to remain silent."

These core principles of polity and jurisprudence should be quite familiar to Anglosphere citizens; if one did not learn them in civics class, they are easily gleaned from TV police dramas. What is not often taught in the schools nor espoused in pop culture, however, is how and why these principles matter for a nation's economic performance. Hannan does an able job of relating arguments made on this point by generations of economists, namely that the stability of property rights that Anglosphere legal and political systems generate is crucial not just for the day-to-day functioning of markets, but for the investments in long-lived capital goods and market innovation that generate sustained economic growth. I will revisit this theme below after addressing the heart of *Inventing Freedom* – a richly detailed account of the historical development of the Anglosphere constitution.

THE ANCIENT ROOTS OF THE ANGLOSPHERE TRADITION

Hannan's major theme is that England, and by extension the Anglosphere, has a unique legal and political institutional structure that is ancient, having emerged from the mists of Germanic and Nordic precursor civilizations, and had already taken on a complete, coherent form by the time of the watershed events of 1066 and 1215. Although the latter date is generally associated with the first *documented* limit upon royal power with Magna Carta, Hannan argues that by this time "the rights of Englishmen" were so well-established that they were already considered ancient.

While the system of English rights and liberties, along with its unique common law-based courts, was well-established by the Middle Ages, it was refined and articulated by

renowned English jurists such as Coke and Blackstone in the 17th and 18th centuries, when it began to take on the form that is still familiar to residents of common law countries to this day.³ The ascendance of this common law tradition, and the individual rights *cum* limitations on governmental power that it entailed, engendered a conflict between the libertarian and absolutist elements in Anglo-American society. This conflict came to a head in what Hannan refers to as the “Anglosphere civil wars.”

The first of these – the English Civil War – culminated in the Glorious Revolution, which Hannan describes as “the moment when the Anglosphere took off, developing into a small-government, individualist, mercantile state-system” (p. 193). The Glorious Revolution also brought on its heels the English Bill of Rights, which in its establishment of individual rights and limitations on government powers, anticipated the ultimate refinement of the Anglosphere constitution which emerged from the American Revolution (which Hannan labels the “Second Anglosphere Civil War”) in the form of the Declaration of Independence and the US Constitution. Noting that the literal sense of the word revolution indicates “a full turn of the wheel,” Hannan suggests that the revolutions of both 1689 and 1776 should not be seen as advances away from an ancient, repressive oligarchy and towards a bright new free and equal future, but indeed as a restoration of the *ancient* traditions of freedom under the law.

CASE STUDY: ANGLOSPHERE DOMINANCE

The Anglosphere civil wars pitted Whigs versus Tories, or, broadly speaking, supporters of the maximal individual liberty and the rule of law versus supporters of royal prerogative and absolute state authority. The Whigs came out on top in both contests, which, as Hannan argues, set the path for the stellar economic growth and geopolitical dominance of the British Empire, followed by the United States, in the 19th and 20th centuries. This success is not surprising in light of the core economic claim about the importance of institutional structure.

Hannan applies this idea broadly to the Anglosphere countries of today, which are defined as those that inherited their legal and political institutions from Mother England.⁴

3 The *Petition of Right*, for example, which was written under Coke’s leadership and presented by Parliament to King Charles I in 1628, is widely seen as the predecessor to the 3rd, 5th, 6th, and 7th amendments to the US Constitution.

4 Hannan lists Australia, Canada, New Zealand, the United Kingdom, and the United States as the core Anglosphere nations; broader definitions also include Ireland, Singapore, Hong Kong, South Africa; some even stretch the concept to encompass India. It is important to note that, although these countries are largely English-speaking, ethnicity is neither a feature nor a requirement of Anglosphere identity. As Hannan notes repeatedly in the book, the Anglosphere has been able to grow geographically and encompass many different nationalities precisely because its core values are built on political ideas about the nature of man, law, and government, rather than on ethnic identity.

Owing precisely to their institutional heritage, they are, not surprisingly, among all nations economically the wealthiest and politically the most free. One does not need to compare nations' real per capita GDP figures to realize this. As Adam Smith had already noticed by 1776, the standard of living in Anglosphere countries is remarkably higher than the rest of the world, including both the 3rd world and the “developed” nations of most of Europe. Again emphasizing the linkage between institutions and development, Hannan cites the consistent ranking of Anglosphere countries among the top of economic freedom indices, such as that published by the Heritage Foundation.⁵ What does this mean in practical terms for entrepreneurs? As Hannan states, “Above all, common law has proved the surest defense of property rights. Today, companies from all over the world pay premiums to sign their contracts in common-law jurisdictions. They do so because they have confidence in the impartiality, security, and fairness of the system” (p. 330).

HANNAN AND THE ECONOMISTS

While Hannan's focus, and the major contribution of this book, is on the historical development of the Anglosphere constitution, his work meshes very well with the stories being told by many economists. Acemoglu, Johnson, and Robinson (2001) have made a compelling empirical case for the importance of institutional structure in economic development, and Acemoglu and Robinson more recently produced a sweeping historical narrative account of how institutional differences, in many cases stretching far back into the mists of time, have had a lasting impact on the economic performance observed in societies today (2012). Acemoglu and Robinson sort nations into two groups: those with “inclusive” institutions that allow widespread participation in political and economic realms, and those with “extractive” institutions that allowed a narrow set of elites to exploit a politically and economically powerless peasant or slave class.

North, Wallis, and Weingast (2009; henceforth NWW) refined and augmented this idea with specific focus on the role of institutional change in economic development. They develop a framework featuring two polar types of polity: the “limited access society” or natural state, and the “open access society,” each defined in terms of its institutional structure. Natural states vest sovereignty among a narrow set of elites, who enjoy economic rents solely on the basis of their elite status. While natural states allowed human societies to eclipse the violence and uncertainty of tribalism, and thus allowed for some stability and economic growth, it took the advent of open access societies to fertilize the industrial and technological progress that has characterized the past several centuries in the West. Open access societies, as the name implies, allow a much broader participation in the political

5 <http://www.heritage.org/index/>

and economic realm, thus opening up peaceful economic competition⁶ which pushes down rents and spurs innovation as main wealth-generating activity.

The inclusive societies of Acemoglu and Robinson, as well as the open access societies described by NWW, feature just those norms that Hannan sets out as the core Anglosphere principles: individual sovereignty, rule of law, and limited government. Thus as the Anglosphere epitomizes the institutional structures requisite for greatest economic success, the enviable performance of the Anglosphere nations on this count should come as no surprise.

This renewed focus on institutions in the economics literature is a welcome development. But it is not a new discovery by any means; one can find this institutional focus at least as far back as Adam Smith: “Little else is requisite to carry a state to the highest degree of opulence from the lowest barbarism, but peace, easy taxes, and a tolerable administration of justice; all the rest being brought about by the natural course of things.” (Stewart 1829, p. 64). A renewed focus came with Mises and Hayek in the 1920s and 30s in the setting of the socialist calculation debate. The Austrians argued that, even assuming a truly benevolent dictator, socialism was doomed in practice due to its institutional shortcomings. They pointed out that abolishing private property in the means of production would eliminate market exchange and thus market prices – the main tool used by businessmen to calculate profit and loss expectations upon which all economic decisions are ultimately based. Thus, socialism would destroy the institutional basis of the market economy.

Another Austrian theme that fits well with Hannan’s hypotheses is the idea that the knowledge essential for the functioning of an advanced economy with widespread division of labor is dispersed – not available to any centralized authority. Hayek argued that decentralized institutions like markets can marshal and organize this dispersed knowledge in an economically coherent manner through a concise communication system of price signals (1948). In contrast to sluggish and punctilious centralized bureaucracies, markets are flexible and adaptable to constantly changing local circumstances, and thus allow for innovation and progress in the economic system. In like manner, the knowledge required for a proper administration of justice⁷ is dispersed, and thus justice is best served under a decentralized, organic, bottom-up legal system – exactly what the Anglosphere gets with its common law tradition.

A related linkage can be seen between Hannan and modern Hayekians in their argument that decentralized institutions, whether economic (e.g. competitive markets) or political (e.g. localized law courts), are “robust,” or resilient against destabilizing shocks. Boettke and Leeson (2004), for example, note that robust institutional structures function well

6 As opposed to predatory competition. Mises highlights this distinction between peaceful, economic competition, which he terms “catallactic competition,” and the violent struggle over a fixed pool of resources, which he labels “biological competition” (1998, p. 274). For competition of the catallactic sort to prevail, what’s needed is something the animal kingdom lacks, but fortunately man can discover and apply: a legal and political structure that recognizes the autonomy and dignity of each individual – in other words, sound *institutions*.

7 Note here that the functioning of the market is built upon a sound legal system, as Smith noted in 1755 and of course as Hannan stresses in *Inventing Freedom*.

in general by tapping dispersed knowledge and engaging many competing, overlapping agents in widespread networks of production and exchange. Thus, if a negative shock – such as natural disaster, honest human error, or outright fraud and rapaciousness – should strike and even destroy several nodes of a network, the remaining nodes and linkages in the system can still marshal the requisite knowledge and resources to maintain some level of productivity or effectiveness. A competitive, decentralized, overlapping production and distribution process for, say, coffee, will ensure that every market and café is continually and reliably supplied on a daily basis, despite any number of afflictions – bad weather in the tropics, shipping delays, human error transcribing and processing orders, fraudulent activities by a handful of dishonest agents, etc. – that would stymie the efforts of a centralized bureaucracy. In like manner, a decentralized legal system, while it might occasionally miscarry justice, will *generally* be able to recognize pertinent circumstances of time and place applicable to a particular case, and see to it that justice is, therefore, *generally* served well. Although Hannan does not apply the particular “robustness” jargon, he acknowledges this basic concept in pointing out the difference between a top-down and bottom-up constitutional structure: “the contracting out of human rights to a charter, necessarily interpreted by some state-appointed tribunal, left the defense of freedom in a small number of hands. If those hands failed, freedom failed. In the Anglosphere, where the defense of freedom was everyone’s business, dictatorship and revolution were almost unknown” (p. 332).

CONCLUSION

In all, Hannan’s book is of immense interest to students of political economy; Hannan lays out a compelling story about the fundamental role of institutions, and particularly constitutional structure, in economic development and the realization of the ideal of a free, open society. For economists and political scientists who adhere to the notion that institutions matter, Hannan’s book provides a rich historical case study worthy of their attention; for Hannan, institutions not only matter, but are the crucial element in national development, both of a flourishing, prosperous economy, and a polity of dignified, autonomous, free people.

Tyler Watts

*Assistant Professor, Economics, Fred Hale School of Business – East Texas Baptist University,
Marshall, TX.*

twatts@etbu.edu

REFERENCES

- Acemoglu, Daron, Simon Johnson and James A. Robinson (2001) 'The Colonial Origins of Comparative Development: An Empirical Investigation.' *American Economic Review* 91 (5): 1369–1401.
- Acemoglu, Daron and James A. Robinson (2012) *Why Nations Fail: The Origins of Power, Prosperity, and Poverty*. Crown Business.
- Boettke, Peter J. and Peter T. Leeson (2004) 'Liberalism, Socialism, and Robust Political Economy.' *Journal of Markets and Morality* 7 (1): 99–111.
- Hannan, Daniel (2010) *The New Road to Serfdom: A Letter of Warning to America*. Harper.
- Hayek, Friedrich A. (1948) *Individualism and Economic Order*. University of Chicago Press.
- Mises, Ludwig von (1998) *Human Action*. Ludwig von Mises Institute.
- North, Douglass C., John Joseph Wallis and Barry R. Weingast (2009) *Violence and Social Orders: A Conceptual Framework for Interpreting Recorded Human History*. Cambridge University Press.
- Stewart, Dugald (1829) *The Works of Dugald Stewart (Vol. VII)*. Cambridge: Hilliard and Brown.

LIST OF ARTICLES

Jan Winiecki: Lessons From Developmental Strategies Over The Last Hundred Years	5
Daniil Gorbatenko: Austrian Business Cycle Theory Without Unrealistic Constructs	12
Jan Winiecki: Lessons From Developmental Strategies Over The Last Hundred Years	29
BOOK REVIEW	
Tyler Watts: Inventing Freedom: How the English-Speaking Peoples Made the Modern World by Daniel Hannan, 2013. Broadside Books, 416 p	49